Minnesota Economics Research Reports

Persuasion and Limited Communication

by

Itai Sher
University of Minnesota

Research Report No. 2008-02, April 2008

Department of Economics

UNIVERSITY OF MINNESOTA
MINNEAPOLIS, MINNESOTA 55455

Persuasion and Limited Communication

Itai Sher[*]
University of Minnesota

This Version: February 2009

Abstract

This paper studies *optimal* persuasion. A speaker must decide which arguments to present and a listener which arguments to accept. Communication is *limited* in that the arguments available to the speaker depend on her information. Optimality is assessed from the listener's perspective assuming that the listener can commit to a *persuasion rule*. I show that this seemingly simple scenario–introduced by Glazer and Rubinstein (2006)–is computationally intractable (formally, NP-hard). However under the assumption known as *normality*, which validates the *revelation principle* in mechanism design environments with evidence (Green and Laffont 1986, Bull and Watson 2007), I show that the persuasion problem reduces to a classic optimization problem, leading to a simple procedure for its solution. This procedure finds not only the optimal rule, but also the *credible implementation* of the optimal rule, i.e., the equilibrium of the game without commitment leading to the same outcome as the optimal rule. Normality also has qualitative consequences for the optimal rule. In particular, under normality, there always exists an optimal rule which is *symmetric*: i.e., *ex ante* equivalent evidence is treated equivalently. When normality fails, all optimal rules may be asymmetric; in other words, the listener may categorize evidence in an arbitrary manner, and base his decisions on these categories in order to influence the speaker's reporting behavior.

JEL Classification: C61, D82, D83.

Keywords: communication, optimal persuasion rules, credibility, commitment, evidence, maximum flow problem.

[*]I am very grateful to Eddie Dekel for many valuable comments and advice. I would also like to thank David Austen-Smith, Marciano Siniscalchi, Asher Wolinsky, David Rahman, Ariel Rubinstein, Jesse Bull, Elchanan Ben-Porath, Melissa Koenig, Peter Sher, Tomasz Strzalecki, and Jan Werner. I am grateful to Evsen Turkay for a conversation that was useful for the proof in Section 4.2, and to Rakesh Vohra for suggesting the relation of the persuasion problem to the set cover problem, as well as an alternative integer program formulation of the persuasion problem. I am also grateful to seminar participants at The University of British Columbia, University of Illinois at Urbana-Champaign, Princeton University, University of Minnesota, University of Rochester, Penn State, Lund University, and Université de Montréal for their comments. All errors are mine.

Contents

1 Introduction

Persuasion is important in a wide array of political and economic environments. People often argue over the best course of action, or attempt to persuade others that their own personal desires are really best for everyone. But what makes an argument persuasive? This paper develops a model that separates persuasive arguments from unpersuasive arguments.

Consider a simple game played between a speaker and a listener. The speaker has some information x which the listener does not have, and would like to persuade the listener to accept a request. There is a set of information states A such that the listener would like to *accept* the request if and only if x belongs to A, and would prefer to *reject* the request otherwise. The speaker can present an argument, but communication is *limited*. In particular, the available arguments depend on the speaker's information x. $\sigma(x)$ is the set of available arguments when the speaker's information is x. There are many reasons why communication may be limited:

1. Arguments may come in the form of hard evidence, or alternatively, the speaker may be required to support her arguments with hard evidence; the available evidence may depend on x.

2. The speaker may be willing to omit information but not to lie (either because of potential penalties if caught, or because of conscience). Alternatively, the speaker may be willing to exaggerate but not too much; she may be willing to tell small lies but not large ones.

3. What a speaker says may depend on what she knows. For example it may be impossible to relate a mathematical proof or a scientific theory if one does not know it.

4. Communication may be subject to time constraints or cognitive and attentional constraints imposed by either the speaker or the listener. For example, the speaker may be allowed only 15 minutes to give a presentation, and even if she has many arguments supporting her position, this may force her to present only one.

Let M be the set of all possible arguments, whose availability depends on the state. A *persuasion rule* is the subset K of the set of arguments M which the speaker finds persuasive. Suppose that the listener commits to a persuasion rule and the speaker knows the persuasion rule K chosen by the listener. For every state x, the speaker presents a persuasive argument (i.e., an argument m in K) if one is available (formally, if $\sigma(x) \cap K \neq \emptyset$); otherwise, the speaker has no choice but to present an unpersuasive argument. A persuasion rule *induces an error* at a state x if the listener would like to accept at x, but ends up rejecting (formally, x is in A but $\sigma(x) \cap K = \emptyset$), or if the listener would like to reject but ends up accepting (formally, x is not in A, but $\sigma(x) \cap K \neq \emptyset$). The *error probability* of a rule is the probability that the rule induces an error. Of course, calculation of the error probability depends on an (exogenous) probability distribution over states. The listener's optimization problem is then as follows:

Persuasion Problem Choose a persuasion rule $K \subseteq M$ that minimizes error probability.

A solution to the persuasion problem therefore induces a division of arguments into two sets: those which are persuasive and those which are unpersuasive. The determination of the persuasiveness of messages is thus reduced to an optimization problem.

In analyzing the persuasion problem, the goal of this paper is to develop a procedure, or *algorithm* for solving the problem. Glazer and Rubinstein (2006) gave a rigorous formulation of the persuasion problem, provided a heuristic method for solving the problem, and solved several interesting instances of the problem. However, Glazer and Rubinstein (2006) did not explicitly analyze an algorithm for persuasion.[1] There is a good reason for this. I show that the persuasion problem is NP-hard (Section 5). This is the first main contribution of this paper. *NP-hardness* is a concept from computer science which, when applied to an optimization problem, is interpreted to mean that the problem is computationally intractable.[2] The notion that the listener decides whether to find an argument persuasive by solving an optimization problem loses some of its plausibility if the optimization problem is difficult to solve. Indeed, if the problem is sufficiently difficult, then it will also be difficult for people to solve it over time through learning or trial and error. Considerations of computational complexity are relevant not only when analyzing persuasion, but in explaining any social interaction as the result of the solution of an optimization problem, or as an equilibrium of some process.

The NP-hardness result may suggest that the goal of finding a reasonable algorithm to solve the persuasion problem is unattainable. However, I identify a sufficient condition for the problem to be computationally tractable. In particular, I show that the problem is tractable in environments where the presentation of evidence satisfies an assumption known as *normality*, which validates a version of the *revelation principle* in mechanism design

[1] Section 5 of Glazer and Rubinstein (2006) is entitled "A Procedure for Finding an Optimal Persuasion Rule". However, this section does not contain an explicit discussion of an algorithm, but rather formulates the persuasion problem as an integer program. The authors begin the section with the sentence "We are now able to prove a proposition that reduces the task of finding an optimal persuasion rule to a simple optimization problem". The problem is simple in the sense that it is simple to state, but it is not simple in the sense that it is simple to solve. Indeed, I show that the persuasion problem is NP-hard. Moreover, there are no known efficient algorithms for solving for solving integer programming in general. While some special types of integer programs are computationally tractable, the general problem of integer programming is NP-hard.

[2] The notion of NP-hardness is part of the theory of computational complexity. There are many references which describe this theory. Several sources which treat the theory at different levels of mathematical difficulty include Garey and Johnson (1979), Papadimitriou (1994), Sipser (2006), and Dasgupta, Papadimitriou, and Vazirani (2008).

environments with evidence (Green and Laffont 1986, Bull and Watson 2007).[3] Formally, normality means that whenever the speaker has several pieces of evidence, there is an action available to her which corresponds to sending all of the evidence. Normality could be violated if the speaker potentially has two pieces of evidence but only has time to present one. Therefore, it has been interpreted as a "no time constraints" assumption.

I show that under normality, the persuasion problem becomes a special case of a classical optimization problem known as the *maximum flow problem*. The maximum flow problem has been intensively studied and there exist many algorithms for its solution. I show how any of these algorithms can be used to find an optimal persuasion rule. This method is the second main contribution of the paper.

Glazer and Rubinstein (2006) presented a striking result about the persuasion problem: commitment has no value in the persuasion problem.[4] In other words, for any optimal persuasion rule, there exists an equilibrium in the game without commitment–a game in which the speaker moves first, and the listener does not commit to a response–which implements the same outcome as the optimal rule, in the sense that the same arguments are found persuasive. I show that not only can any algorithm for the maximum flow problem be used to find an optimal persuasion rule in normal persuasion problems, but any such algorithm can also be used to find the *credible implementation* of the optimal rule–the equilibrium of the game without commitment implementing the same outcome as the optimal rule. So it is not more *computationally* difficult for the speaker and listener to settle on an equilibrium implementing the optimal rule than it is for the listener to find an optimal rule.

The relationship between the persuasion problem and the maximum flow problem also leads to several qualitative conclusions about optimal rules under normality. Glazer and Rubinstein (2006) present examples in which optimal persuasion rules treat differently two pieces of evidence which are essentially equivalent in terms of their content. For instance, if the evidence consists of the written testimony of certain experts, an optimal persuasion rule may take into account the gender of the experts even though this really has no bearing on the matter at hand. Moreover, *all* optimal rules in this example must condition acceptance of the speaker's request on some arbitrary category such as gender. I establish that under normality, there always exists an optimal persuasion rule which is symmetric in the sense that it treats equivalent evidence equivalently. So it is only in the presence of time constraints that optimal persuasion rules must treat equivalent evidence differently. I also establish monotone comparative statics results concerning how the difficulty of persuading

[3] Green and Laffont (1986) referred to a related condition known as the *nested range condition*. Along with these papers, several other papers including Lipman and Seppi (1995), Singh and Wittman (2001), Bull and Watson (2004), Forges and Koessler (2005), Deneckere and Severenov (2008), Ben-Porath and Lipman (2008), and Kartik and Tercieux (2009) have explored the consequences of related properties in mechanism design environments where agents possess hard evidence.

[4] Similar results can be found in Glazer and Rubinstein (2001) and Glazer and Rubinstein (2004) in the context of related but distinct models.

the listener and the probability that the listener selects the wrong action change as the ex ante probability that the listener would like to accept the speaker's request changes. Again, the normality assumption is critical. From a technical perspective, the qualitative results concerning symmetry and comparative statics derived under the assumption of normality stem from the fact that the maximum flow problem is an instance of *supermodular optimization*, and the general properties of supermodular optimization problems.

It is well known that in general, the analysis of properties of Nash equilibrium is more computationally intractable than the analysis of correlated equilibrium (Gilboa and Zemel 1989). The credibility result implies that finding the optimal persuasion rule is equivalent to finding the best Nash equilibrium (also, the best sequential equilibrium) of the persuasion game without commitment. The difference between Nash equilibrium and correlated equilibrium (or communication equilibrium) is that in the latter case the game is augmented with a mediator and additional communication possibilities. This suggests that augmenting the persuasion problem with additional communication *possibilities* might generally make the problem more computationally tractable. However, the model of persuasion is explicitly concerned with communication *limitations*, so such an augmentation requires interpretation. These issues are discussed in Section 8.

2 Examples

This section will present three canonical examples from the literature on persuasion games. These examples demonstrate interesting properties of persuasion. In all the examples, the speaker's preferences are common knowledge, but the messages that the speaker can send are known only to the speaker and depend on the state of the world. Unlike some of the examples below, this paper focuses on situations in which the listener faces a binary decision because the main ideas which I focus on can be analyzed in this case. While there are some subtleties,[5] many of the results can be extended to situations where the listener's decision is not binary. Sher (2008) studies the persuasion problem with multiple actions.

Example 1 This example is from Milgrom (1981) and Milgrom and Roberts (1986), and is similar to Grossman and Hart (1980) and Grossman (1981). A seller has a product whose quality $x \in X := \{0, 1, \ldots, n\}$ is privately known to the seller. $p_x > 0$ is the probability that the quality is x. When the quality is x, the seller may tell a buyer that the quality is in any set $Y \subseteq X$ with $x \in Y$. So, the seller may withhold information but may not lie. The communication constraint that $x \in Y$ allows us to interpret any statement Y as hard evidence that the state is in Y. After observing the seller's report, the buyer may purchase any quantity $q \in \mathbb{R}_+$ at price π, which is exogenous. The buyer's utility is given

[5]The main issue is that the credibility result does not hold in general with multiple actions. However, Sher (2008) finds conditions under which it does hold.

by $v(x, q) - \pi q$. Assume that v is continuously differentiable and strictly concave in q, that $\frac{\partial}{\partial q} v(0, 0) > 0$ and that $\frac{\partial}{\partial q} v(x, q)$ is strictly increasing in x. The seller's utility is given by $(\pi - c)q$ where c is the unit cost. Milgrom (1981) and Milgrom and Roberts (1986) show that in every sequential equilibrium of this game (i) the buyer always infers the true quality, and hence purchases the optimal quantity given the quality, and (ii) given any message Y, the speaker has *skeptical* beliefs in that he assumes that the quality is the lowest quality in Y. This is known as the *unraveling result* in the persuasion game literature. The result has been used to argue that with sufficient availability of evidence, laws mandating the disclosure of evidence would have no effect because even without such laws, there would be full disclosure in equilibrium.

Example 2 This example is from Shin (2003), and is similar to Shin (1994a), Shin (1994b) and Dziuda (2007). A firm undertakes N projects, each of which succeeds with probability r, which is independent across projects. In period 1, each project is completed with probability θ, which is independent across projects. The firm's manager privately knows which of the completed projects have been successful but knows nothing about uncompleted projects. A state is a pair (S, F) where S (respectively, F) is the number of successful (respectively, unsuccessful) projects at period 1. The manager may report any pair (S', F') with $S' \leq S$ and $F' \leq F$ to the market. S' (respectively, F') is the number of *reported* successes (respectively, failures). So, the manager may withhold information, but may not lie. At period 2 the remaining projects are completed and the outcomes of all projects are revealed to the market. The firm then has a value of $W = B^{S^*} C^{N-S^*}$, where $0 < C < B$ and S^* is the total number of successes over projects which were completed in either period. The market is treated as a player who chooses a price π at period 1 after the manager's report, and gets utility $-(\pi - W)^2$. So the market would like to choose a price as close to the value as possible and in particular, to minimize the square loss. The manager's objective is to maximize π. Shin shows that there is a sequential equilibrium in which the manager uses a *sanitization* strategy whereby he reveals all successes and no failures. There does not exist an equilibrium with full disclosure. The difference between this example and the previous one is that here there is uncertainty how much information the speaker has, which mutes the listener's skepticism.

Example 3 This example is from Glazer and Rubinstein (2006) and is similar to examples in Fishman and Hagerty (1990), Glazer and Rubinstein (2001), and Glazer and Rubinstein (2004). A speaker attempts to persuade a listener that the majority of evidence supports the speaker's opinion. The speaker observes five facts, each of which may either support or oppose the speaker's position. This is represented as a state $x \in \{0, 1\}^5$, where $x_i = 0$ means that fact i opposes the speaker's position, and $x_i = 1$ means that fact i supports the speaker's position. The speaker only has time to show the listener 2 facts. So a message can be represented as an element $m \in \{-1, 0, 1\}^5$, where $m_i = -1$ means that the ith fact has been omitted. At state x, the listener may present any message m such that $m_i = -1$ for at least three indices i and otherwise $m_i = x_i$. Again, the speaker cannot lie. However, in contrast to the previous examples, the speaker *must* withhold some of his information. The speaker makes a request of the listener and the listener's goal is to accept the speaker's request if and only if the majority of the facts support the speaker's position. Say that the listener makes an error in any state where he fails to achieve his goal. Suppose that each state $x \in X$ occurs with probability 1/32. The listener's objective is to minimize the probability of error. Suppose that the listener can commit to how he will respond to any message. Then it is optimal for the listener to partition the indices $\{1, 2, 3, 4, 5\}$ into two sets of size 2 and 3, which we will call *categories*, for example $\{1, 2\}$ and $\{3, 4, 5\}$, and to accept the speaker's request only if the speaker presents two facts supporting his opinion *in the same category*. For example, if the evidence consists of the opinions of five experts, two of which are women and three of which are men, it would be optimal to require the speaker to present supporting opinions of two experts *of the same gender* in order to win acceptance. Categories such as gender are ex ante irrelevant but the optimal rule may have to make use of some such categories. Glazer and Rubinstein (2006) interpret this as a pragmatic phenomenon, where pragmatics is the subfield of linguistics that studies the nonliteral meanings of words which arise in conversation. The authors view the persuasion model is contributing to pragmatics in noncooperative settings. Observe that since even under commitment, there is a positive probability of error, there will not be full information revelation in the game without commitment, in contrast to Example 1.

3 The Persuasion Problem

Glazer and Rubinstein (2006) present the following problem. Because the formulation allows for persuasion rules to employ randomization, it is more general than the formulation presented in the introduction. A speaker would like to persuade a listener to accept a request. There is a finite set of states X and a subset A of X such that the speaker would like to accept the request if and only if the state is in A. Define $R := X \setminus A$. For any $x \in X$, let p_x be the probability of x. Assume that for all $x, p_x > 0$. Let $p = (p_x)_{x \in X}$. For every state $x \in X$, there is a finite set $\sigma(x)$ of statements which are available at x. Let $M = \bigcup_{x \in X} \sigma(x)$.[6] We refer to (X, M, σ) as the **message structure** and to σ as the **message correspondence**. A **persuasion problem** is a tuple (X, M, σ, A, p).

The listener may select a **persuasion rule** f which is a function $f : M \to [0, 1]$. $f(m)$ is the probability that the listener accepts the speaker's request conditional on statement m. A persuasion rule is **deterministic** if for all m, $f(m) \in \{0, 1\}$. Let F be the set of all

[6] We also allow $M \supseteq \bigcup_{x \in X} \sigma(x)$.

persuasion rules. Given a state x, define $\alpha(f,x) := \max_{m\in\sigma(x)} f(m)$. This is the maximal probability of acceptance that a speaker can induce given the persuasion rule f at state x. Define $\alpha(f) := (\alpha(f,x) : x \in X)$. Define $\mu_x(f) := 1-\alpha(f,x)$ if $x \in A$, and $\mu_x(f) := \alpha(f,x)$ if $x \in R$. $\mu_x(f)$ is the probability of error induced by persuasion rule f in state x. The listener would like to solve: $\min_{f\in F} \sum_{x\in X} p_x\mu_x(f)$. The objective is to minimize the **error probability**. The formulation of the problem is such that the listener may commit to a persuasion rule, and the speaker knows of this commitment at the time that he selects this message.

Glazer and Rubinstein (2006) prove that there always exists an optimal persuasion rule which is deterministic; the more simplified description of the persuasion problem in the introduction depends on this result. Glazer and Rubinstein (2006) also provide an integer program whose solutions correspond to the state by state error probabilities induced by optimal persuasion rules. Define an L to be a pair (x,T) such that $T \subseteq R$ and $x \in A$, and $\sigma(x) \subseteq \bigcup_{y\in T} \sigma(y)$. An L, (x,T) is **minimal** if for all proper subsets T' of T, (x,T') is not an L.

Theorem 1 *(Glazer and Rubinstein (2006)) Let $(\mu^*_x)_{x\in X}$ be a solution to:*

$$\begin{aligned} & \min_{\{\mu_x\}_{x\in X}} \sum_{x\in X} p_x\mu_x \\ \text{s.t.} \quad & \mu_x \in \{0,1\}, \ \forall x \in X \\ & \sum_{y\in\{x\}\cup T} \mu_y \geq 1 \ \text{ for every minimal } L,\ (x,T). \end{aligned} \tag{1}$$

Then there is an optimal persuasion rule that induces error probabilities $(\mu^*_x)_{x\in X}$. Moreover, any optimal deterministic persuasion rule induces a vector of error probabilities equal to some solution $(\mu^*_x)_{x\in X}$ of (1).

The linear constraints in (1) are referred to by the authors as the **L-principle**, and are perhaps more easily understood, if rewritten as: $\sum_{y\in T} \mu_y \geq 1-\mu_x$ for every L, (x,T). Suppose that the listener would like to avoid a mistake in state $x \in A$. Then the listener must accept the listener's request in state x, so there must be some message $m \in \sigma(x)$ that the listener accepts. Then $\mu_x = 0$, so $\sum_{y\in T} \mu_y \geq 1$. Why? Because at least one member y of T has message m as well, and therefore the listener would have to accept the request at state y as well, making a mistake, so that $\mu_y = 1$.

Nothing essential in the above model changes if we introduce a state dependent error cost $\ell_x > 0$. The listener's objective is then $\min_{f\in F} \sum_{x\in X} \ell_x p_x\mu_x(f)$. Theorem 1 remains valid if the objective is changed to $\min_{\{\mu_x\}_{x\in X}} \sum_{x\in X} \ell_x p_x\mu_x$.

3.1 The Game Without Commitment

In the **game without commitment**, the speaker first sends a message and the listener responds with either acceptance or rejection. The speaker's strategy is a function $\zeta : X \times M \to [0,1]$, where $\zeta(x,m)$ is interpreted as the probability that the speaker plays m at state x. It is assumed that $\zeta(x,m) > 0$ only if $m \in \sigma(x)$ since only available messages can be sent with positive probability. The listener's strategy can still be represented by a function $f : M \to [0,1]$. Glazer and Rubinstein (2006) establish that if f is an optimal persuasion rule, then there exists a speaker strategy ζ such that (ζ, f) is a Bayesian Nash equilibrium of the game without commitment. Thus commitment has no value in this problem. I refer to this as the **credibility result**, and to the equilibrium (ζ, f) as the **credible implementation**.

A possible objection concerns the use of Bayesion Nash equilibrium rather than sequential equilibrium in the credibility result. The objector may fear that without appeal to sequential equilibrium, it is possible that the equilibrium may be supported by threats off the equilibrium path which are not "credible". The objection is answered by the following theorem:

Theorem 2 *Assume that persuasion rule f and speaker strategy ζ satisfy the following properties:*

i f is an optimal persuasion rule.

ii (ζ, f) is a Bayesian Nash equilibrium in the game without commitment.

Then there exists a persuasion rule f' such that:

1. *f' is an optimal persuasion rule.*
2. *$\alpha(f) = \alpha(f')$.*
3. *(ζ, f') is a sequential equilibrium in the game without commitment.*

Proof. In Appendix. □

Property 2 of the theorem says that f and f' are outcome equivalent in the sense that they induce the same mappings from states to probabilities of acceptance. Theorem 2 and the credibility result together imply that every optimal persuasion rule is outcome equivalent to an optimal persuasion rule which can be implemented as a sequential equilibrium in the game without commitment.

4 Two Simple Proofs of the Credibility Result

In this section, I present two new simple proofs of the credibility result. The purpose of this section is to provide a clear explanation of why the credibility result is true. The first proof is based on the the minimax theorem. The proof constructs a zero-sum game which is related to the persuasion game, and then derives the credibility result from the fact that commitment has no value in a zero-sum game. The second proof is in a sense more elementary, although it uses the existence of Nash equilibrium, which is a generalization of the minimax theorem.

4.1 Proof Based on the Minimax Theorem

Fix an optimal persuasion rule f^* in a persuasion problem $Q = (X, M, \sigma, A, p)$. For each $x \in A$, let $m^x \in \text{argmax}_{m \in \sigma(x)} f(m)$. Define $\sigma^*(x) := \{m^x\}$ if $x \in A$ and $\sigma^*(x) := \sigma(x)$ if $x \in R$. Consider the persuasion problem $Q^* := (X, M, \sigma^*, A, p)$, and let Γ^* be the corresponding game without commitment. Let f be any persuasion rule. At every state in R, f induces the same acceptance probability in Q^* as in Q. In every state in A, f induces a weakly lower acceptance probability in problem Q^* than in Q. It follows that every persuasion rule induces a weakly higher error probability in Q^* than in Q. However f^*–which is optimal in Q–induces the same error probability in Q^* as in Q. So f^* is optimal in Q^*. Let $U(\zeta, f)$ be the error probability given strategy profile (ζ, f) in Γ^*.[7] So the listener chooses f to minimize $U(\zeta, f)$. Assuming that the speaker's objective is to choose ζ to maximize $U(\zeta, f)$ does not alter the speaker's best response correspondence in Γ^* since the speaker is only free to make a nontrivial choice in R. That f^* is optimal in Q^* means that f^* is a minimax strategy for the listener in Γ^*, and hence there is an equilibrium (ζ^*, f^*) in Γ^*. It follows from the fact that $\zeta^*(x, m^x) = 1$ for all $x \in A$ and the definition of m^x that (ζ^*, f^*) is also a Bayesian Nash equilibrium in the game without commitment corresponding to Q.

4.2 Proof Based on the Existence of Nash Equilibrium

This proof applies to deterministic optimal persuasion rules.[8] We know that deterministic optimal rules always exist. Let f^* be a deterministic optimal rule in persuasion problem $Q := (X, M, p, \sigma, A)$. Consider a pair of persuasion problems $Q^i := (X^i, M^i, p^i, \sigma^i, A^i)$ for $i = 0, 1$ in which $X^i := \{x \in X : \alpha(f^*, x) = i\}, M^i := \{m \in M : f^*(m) = i\}, p^i_x := p_x / \sum_{y \in X^i} p_y$, $\sigma^i(x) := \sigma(x) \cap M^i$, $A^i = A \cap X^i$.[9] Note that $\sigma^i(x)$ can never be empty. There exists an equilibrium (ζ^i, f^i) in the game without commitment corresponding to Q^i. Let $K := \{m \in M^1 : f(m) < 1\}$. The listener weakly prefers to reject each message in (ζ^1, f^1). Assume for contradiction that the listener strictly prefers to reject some message $m \in K$ in (ζ^1, f^1). It follows that the speaker strictly prefers to reject conditional on the

[7] Notice that since F already allows randomization, allowing the speaker to randomize over persuasion rules would not change the strategy set.

[8] Similar ideas could be used to extend the proof to optimal persuasion rules that are not deterministic.

[9] If $X^i = \emptyset$, then the proof is even simpler, and does not include Q^i, but focuses on $Q^j = Q$ for $j \in \{0, 1\}$ and $j \neq i$.

state being in the set $\{x \in X^1 : \sigma^1(x) \cap M^1 \setminus K = \emptyset\}$. It follows that the deterministic persuasion rule that accepts exactly the messages in $M^1 \setminus K$ gives the listener a higher expected utility than f^* in Q, a contradiction. A similar argument shows that the listener does not strictly prefer to accept any message in (ζ^0, f^0). Defining ζ^* to agree with ζ^i on X^i for $i = 0, 1$, it follows that (ζ^*, f^*) is an equilibrium in the game without commitment corresponding to Q.

5 The Persuasion Problem is NP-hard

In this section I examine the question of how difficult it is to find an optimal persuasion rule. The formal tools I use to assess the difficulty of the persuasion problem are taken from computer science. A problem is considered to be tractable if the running time required for a Turing machine–a formal model of a general purpose computer–to solve the problem is bounded by a polynomial function of the length of the input necessary to describe the instance of the problem being solved. The class of problems possessing a polynomial time algorithm is known as P. A problem is considered difficult if there is no polynomial time algorithm for solving it. Unfortunately, it has proven very difficult to prove that computational problems are hard in this sense. The approach that has been taken in computer science is as follows. There is a more powerful model of computation known as a *nondeterministic* Turing machine. A nondeterministic Turing machine is not considered to be a reasonable model of a computer because it can take "guesses" during the course of its computation; nevertheless, it is a useful theoretical construct. A computational problem is in the complexity class NP if a nondeterministic Turing machine can solve the problem in polynomial time. It is widely believed–but not known–that there exist problems in NP with a worst case running time which is exponential on an ordinary Turing machine. A problem is NP-hard if it is at least as hard as any problem in NP. A little more formally, a problem Z is NP-hard if for any problem Y in NP there is a polynomial time algorithm that reduces Y to Z. Showing that a problem is NP-hard is accepted as strong evidence that the problem is computationally intractable. For a formal treatment of all of these notions, see Garey and Johnson (1979), Papadimitriou (1994), or Sipser (2006).

Theorem 1–due to Glazer and Rubinstein (2006)–reduced the the persuasion problem to an integer program (1). An integer program is a linear program with additional constraints specifying that the variables must take integer values. Integer programming is NP-hard. However, since (1) is a special case of integer programming it could in principle be more tractable than the general case. For example, if the integer constraints were not binding, then (1) reduces to a linear program;[10] linear programming possesses polynomial time algorithms. However, the following example shows that unfortunately the integer constraints

[10]However, even in this case, (1) may have an exponential number of constraints, which could be another problem.

in (1) may be binding.

Example 4 Suppose that $X = \{x_1, x_2, x_3, y_1, y_2, y_3\}$, $A = \{x_1, x_2, x_3\}$, so that $R = \{y_1, y_2, y_3\}$. Suppose that $p_{x_1} = p_{x_2} = p_{x_3} =: p_x$ and $p_{y_1} = p_{y_2} = p_{y_3} =: p_y$. Suppose further that $p_x > p_y$. Suppose that for all $i \in \{1,2,3\}$ and j,k such that $\{i,j,k\} = \{1,2,3\}$, $\sigma(x_i) = \{m_j, m_k\}$, and $\sigma(y_i) = \{m_i\}$. Then the persuasion problem simplifies to:

$$\min[p_x(\sum_{i=1}^{3} \mu_{x_i}) + p_y(\sum_{i=1}^{3} \mu_{y_i})] \tag{2}$$

$$s.t \quad \mu_{x_i} + \mu_{y_j} + \mu_{y_k} \geq 1, \quad \forall i \in \{1,2,3\}, \forall j,k \text{ with } \{i,j,k\} = \{1,2,3\} \tag{3}$$

$$\mu_z \in \{0,1\}, \quad \forall z \in X. \tag{4}$$

Adding up the constraints of the form (3), we see that:

$$\sum_{i=1}^{3} \mu_{x_i} + 2\sum_{i=1}^{3} \mu_{y_i} \geq 3$$

It is then easy to see invoking (4) that (*) there must exist at least two distinct elements z and z' of X such that $\mu_z = \mu_{z'} = 1$. Now consider the following proposed solution: $\mu_z = 1$ if $z \in \{y_1, y_2\}$, and $\mu_z = 0$ otherwise. It is easy to see that this proposed solution is feasible in (2)-(4). Moreover, it follows from (*) and the fact that $p_x > p_y$, that the proposed solution is optimal, showing that the persuasion problem has value $2p_y$.

Next replace the integer constraints (4) by the relaxed constraints:

$$\forall z \in X, \qquad \mu_z \in [0,1] \tag{5}$$

Then consider the proposed solution $\mu_z = 0$ if $z \in A$, and $\mu_z = 1/2$ if $z \in R$. It is easy to see that this solution is feasible in the relaxed program and attains a value in the relaxed program of $(3/2)p_y$ which is less than the value $2p_y$ which was found above in the unrelaxed program. It follows that the integer constraints (3) are binding. □

Theorem 3 *The persuasion problem is NP-hard.*

Proof. In Appendix. □

Theorem 3 shows that the listener must solve a computationally hard problem in order to find an optimal persuasion rule. Similarly if we think of the credible implementation– i.e., the equilibrium in the game without commitment implementing the optimal rule–as arising through some adaptive or cognitive process, this process must implicitly solve a computationally hard problem. This reduces the plausibility of the notion that the listener would select an optimal rule or that an optimal rule would be used in equilibrium. Equilibria which require less computation–if such equilibria exist–may be more likely to occur than those that implement optimal rules. These considerations motivate the search for special cases of the persuasion problem which require less computation, which is the task we turn to next.

6 Normal Persuasion Problems

In this section, I introduce a property on the message structure called *normality*. This property has an intuitive interpretation and ensures that the persuasion problem is tractable. Normality was introduced by Bull and Watson (2007), and is closely related to the *nested range condition* of Green and Laffont (1986).[11]

Definition 1 *A message structure* (X, M, σ) *is* ***normal*** *if:*

$$\forall x \in X, \exists m_x \in \sigma(x), \forall y \in X, m_x \in \sigma(y) \Rightarrow \sigma(x) \subseteq \sigma(y).$$

m_x *is* x*'s* ***maximal message***.[12]

Intuitively, m_x summarizes all of x's information. If m_x is available at state y, then all messages available at x are also available at y. Normality can be interpreted as the **absence of time constraints** in communication: without time constraints, it is possible to present a message which summarizes all of one's evidence, simply by presenting all of one's evidence; with time constraints, this is impossible. I now categorize the examples in Section 2 with respect to normality.

Example 1 The set of states is $X = \{0, 1, \ldots, n\}$. The set of messages M is the set of all nonempty subsets of X. The message correspondence is $\sigma(x) = \{Y \subseteq X : x \in Y\}$ This is a normal message structure with $m_x = \{x\}$ for all $x \in X$.
Example 2 $X = M = \{(S, F) \in \{0, 1, \ldots, N\}^2 : S + F \leq N\}$. $\sigma(S, F) = \{(S', F') \in M : S' \leq S, F' \leq F\}$ is normal with $m_x = x$ for all $x \in X$.
Example 3 $X = \{0, 1\}^5$. $M = \{m \in \{-1, 0, 1\}^5 : |m_i = -1| \geq 3\}$. The message correspondence is $\sigma(x) = \{m \in M : m_i \neq -1 \Rightarrow m_i = x_i\}$. σ is not normal. To see this, consider $x = (1, 1, 1, 1, 1)$ and $m \in \sigma(x)$. For example, let $m = (1, 1, -1, -1, -1)$. Note that $(1, 1, -1, -1, -1) \in \sigma(1, 1, 1, 1, 0)$ but $\sigma(1, 1, 1, 1, 1) \not\subseteq \sigma(1, 1, 1, 1, 0)$ because $(-1, -1, -1, 1, 1) \notin \sigma(1, 1, 1, 1, 0)$. So m cannot be m_x. A similar argument holds for any $m \in \sigma(x)$.

[11] Several other papers on mechanism design with evidence and related topics note the importance of these notions or similar notions, including Lipman and Seppi (1995), Deneckere and Severinov (2001), Singh and Wittman (2001), and Forges and Koessler (2005).

[12] The maximal message may not be unique. However, if there are two messages m and m' such that both satisfy the definition of x's maximal message, then it must be the case that for all $y \in X$, $m \in \sigma(y) \Leftrightarrow m' \in \sigma(y)$; in other words, the messages are equivalent.

Observe that in the intuitive explanation of Examples 1 and 2 in Section 2, there were no time constraints, and these message structures are normal. The explanation of Example 3 involved time constraints, and indeed it is not normal. In any persuasion problem in which at each state, the speaker receives n pieces of information–where n may even be random– but at least sometimes, has time only to present some but not all of the information, the persuasion problem is not normal.[13] The argument is similar to that presented with regard to Example 3 above.

There is a very rich family of message correspondences which satisfy normality. To see this observe that for any quasi-order (i.e., reflexive and transitive relation) $\preccurlyeq$, and any set X, the message correspondence $\sigma(x) := \{y \in X : y \preccurlyeq x\}$ is normal with $m_x = x$.[14] The interpretation of $x \preccurlyeq y$ is that y can mimic x. There are many variations of the message structure in Example 2 which would still be normal. For example, suppose that rather than presenting the number of favorable arguments and the number of unfavorable arguments, each favorable argument had specific characteristics. The speaker could present evidence that for example, "project 1 was successful and project 3 was successful, while project 5 was a failure", rather than presenting evidence that two project were successful and one was a failure. If the speaker observed a subset of the project outcomes and could present any subset of what he observed, the message structure would still be normal. One could also allow for the possibility that the speaker could observe more than just success or failure, but degrees of success or other characteristics of the outcome. If the speaker has the option of presenting all of his evidence or some parts of it, then the message structure is still normal. In any such example, normality allows for the possibility that like in Example 2, the listener does not know how much evidence the speaker has collected.

The final observation of this section is that it is easy to verify that a message correspondence satisfies normality.

Theorem 4 *There exists a polynomial time algorithm that determines whether a message structure (X, M, σ) is normal, and if so finds the maximal message m_x at each state x.*[15]

[13] The pieces of evidence are assumed to be truly distinct in the sense that there does not exist at every state one piece of evidence from which it is possible to infer the existence and identity of all the others.

[14] A similar observation is made by Green and Laffont (1986).

[15] There are many ways to see that the property is normality is easy to verify. One way is to construct a simple algorithm that does this. Another way is as follows: assume without loss of generality that X and M are disjoint. We can represent a message structure as a directed graph (V, E) with vertices $V = X \cup M$ and edges $E := \{(x, m) : x \in X, m \in M, m \in \sigma(x)\}$. Then normality can be expressed by the following sentence in first-order logic with the vocabulary of graph theory:

$$\forall y, \exists z, [(y, z) \in E \text{ and } [\forall w, [(z, w) \in E \Rightarrow \forall u, [(y, u) \in E \Rightarrow (z, u) \in E]]]]$$

The existence of a polynomial time algorithm now follows from Theorem 5.1 of Papadimitriou (1994), which says that any property of graphs which is expressible by a first-order sentence over the vocabulary for graphs can be decided in polynomial time.

I will briefly explain one way that this can be done. Observe first that if m' is the maximal message at x, then $m' \in H(x) := \text{argmin}\{|\sigma^{-1}(m)| : m \in \sigma(x)\}$. It is easy to calculate $|\sigma^{-1}(m)|$ for each message m, and hence easy to find a message in $H(x)$ for each state x. So, for each state x, pick an arbitrary message $m' \in H(x)$ and check whether $\sigma^{-1}(m') \subseteq \sigma^{-1}(m)$ for all $m \in \sigma(x) \setminus \{m'\}$. If the answer is "yes", then m' satisfies the condition imposed by Definition 1 for the maximal message m_x at x. If the answer is "no" for even one state, then the message structure is not normal.

6.1 Simplifying the Persuasion Problem

The following lemma relates normality to the L's of Glazer and Rubinstein (2006).

Theorem 5 *1. In a normal persuasion problem, every minimal L, (x, T) is such that $T = \{y\}$ for some $y \in R$.*

2. Fix (X, M, σ, p). If for all $A \subseteq X$, every minimal L, (x, T) is such that $T = \{y\}$ for some $y \in R$, then the message structure is normal.

Proof of 1. Consider an L, (x, T). Then $\sigma(x) \subseteq \bigcup_{y \in T} \sigma(y)$. So there exists $y \in T$ such that $m_x \in \sigma(y)$. Then by normality $\sigma(x) \subseteq \sigma(y)$. So $(x, \{y\})$ is an L.
Proof of 2. If σ is not normal, there exists $x \in X$ such that for all $m \in \sigma(x)$, there exists $y_m \in X$ such that $m \in \sigma(y_m)$ but $\sigma(m) \not\subseteq \sigma(y_m)$. Let $A = \{x\}$. Then $(x, \{y_m : m \in \sigma(x)\})$ is an L, but for every $m \in \sigma(x)$, $(x, \{y_m\})$ is not an L. □

Combining Theorem 5 with Theorem 1 yields the following corollary:

Corollary 1 *Suppose that the persuasion problem is normal. Let $(\mu_x^*)_{x \in X}$ be a solution to:*

$$\begin{aligned}
&\min_{\{\mu_x\}_{x \in X}} \sum_{x \in X} p_x \mu_x \\
\text{s.t.} \quad &\mu_x \in \{0, 1\}, \ \forall x \in X \\
&\mu_x + \mu_y \geq 1 \ \text{ for all } x \in A \text{ and } y \in R \text{ such that } \sigma(x) \subseteq \sigma(y).
\end{aligned} \tag{6}$$

Then there is an optimal persuasion rule that induces error probabilities $(\mu_x^*)_{x \in X}$. Moreover, any optimal deterministic persuasion rule induces a vector of error probabilities equal to some solution $(\mu_x^*)_{x \in X}$ of (6).

The integer program (6) corresponds to a well known combinatorial optimization problem known as **minimum weight vertex cover problem on bipartite graphs**.[16] It is well-known, and easy to confirm that the constraint matrix for this problem is totally unimodular,[17] which implies the following corollary:

[16] See Section 12.2 of Hochbaum (2008).
[17] See Section 12.2 of Hochbaum (2008). Note that the qualification "on bipartite graphs" is important here; the minimum weight vertex problem without this qualification is NP-hard. In the program, (6) the

Corollary 2 *Suppose that the persuasion problem is normal. Then the integer constraints are not binding. In other words, replacing the integer constraints of the form* $\mu_x \in \{0, 1\}$ *by constraints of the form* $0 \leq \mu_x \leq 1$ *does not change the value of the problem.*

This result contrasts with Example 4, which showed that in the general formulation of the persuasion problem, it is possible that the integer constraints are binding. Notice also that, under normality, the number of constraints is bounded above by $|X|^2$.[18] Since there is a polynomial time algorithm that solves linear programming, the following theorem follows.

Theorem 6 *There is a polynomial-time algorithm which solves the persuasion problem under the assumption of normality. In contrast, if* $P \neq NP$*, then no such algorithm exists for the persuasion problem without normality.*

I conclude this section by giving an interesting formulation and interpretation of the persuasion problem under normality which follows from Corollary 2. Consider the following problem:

Mimicking Problem Select a collection $Y \subseteq X$ of states to minimize error:

$$\sum_{x \in A \setminus Y} p_x + \sum_{x \in R \cap Y} p_y$$

subject to the constraint that if $x \in A \cap Y$, $y \in R$ and $\sigma(x) \subseteq \sigma(y)$, then $y \in Y$. In other words if $x \in A$ is accepted and $y \in R$ can mimic x, then y is also accepted.

---

key point is that we can separate the elements of X into two subsets A and R, such that the constraint $\mu_x + \mu_y \geq 1$ applies only if $x \in A$ and $y \in R$. The integer program which replaces this constraint by $\mu_x + \mu_y \geq 1$ for all edges $(x, y) \in E$, where E is the set of edges in an arbitrary graph with vertices X is intractable.

[18] The observation about the number of constraints is important because a problem which can be formulated as a linear program whose number constraints grow exponentially in the description of the problem may not admit a polynomial time algorithm. If one looks at the general formulation of the persuasion problem as a linear program (1), then if, for example $|X|$ is odd, there may be as many $\binom{|X|-1}{(|X|-1)/2}$ minimal L's. To see this, suppose that $A = \{x^*\}$ for some $x^* \in X$ and suppose that for each $S \subseteq R$ with $|S| = \frac{|R|}{2} + 1$, there is a message m_S with $m_S \in \sigma(y) \Leftrightarrow y \in S \cup \{x\}$. Then for every $T \subseteq R$ with $|T| = \frac{|R|}{2}$, (x^*, T) is a minimal L so that there are $\binom{|X|-1}{(|X|-1)/2}$ minimal L's. However, the problem with the general persuasion problem is not really that there must be too many constraints in the integer programming formulation, but rather is that the integer constraints are binding. Establishing this, Rakesh Vohra has suggested the following alternative integer program for the persuasion problem which is guaranteed to have polynomially many constraints:

$$\min \sum_{x \in X} p_x \mu_x$$
$$\mu_x + \sum_{m \in \sigma(x)} f(m) \geq 1, \ \forall x \in A$$
$$\mu_x \geq f(m), \ \forall x \in R$$
$$\mu_x \in \{0, 1\}, \ \forall x \in A$$
$$f(m) \in \{0, 1\}, \ \forall m \in M$$

In the mimicking problem the set Y corresponds to the deterministic persuasion rule f which is such that $f(m) = 1$ if $m = m_x$ for some $x \in Y \cap A$ and $f(m) = 0$ otherwise. A simple manipulation of (6) shows that under normality, the persuasion problem is equivalent to the mimicking problem. In general, without normality, the two problems are not equivalent. The mimicking problem is essentially the same as another well-known optimization problem, which is also equivalent to (6) known as the **maximal closure problem** (Picard 1976).

6.2 An Algorithm for Persuasion

The previous section established that there exists a polynomial time algorithm for solving the persuasion problem under normality. In this section, I will show something stronger, namely that under normality, there exists a polynomial time algorithm that finds not only the optimal persuasion rule but also the credible implementation of the optimal rule. This means that it is not more *computationally* difficult for the speaker and listener to settle on an equilibrium implementing the optimal rule than it is for the listener to find an optimal rule. Unlike in the previous section, where I only proved that it is possible to find a polynomial time algorithm for finding an optimal rule, I will discuss in detail one such algorithm to give the reader insight into the nature of a reasoning process that solves the persuasion problem. However, as will be made clear, there is a large class of algorithms that could be used instead of the one that I use.

The algorithm which I employ exploits the fact that (6) is a special case of the well known **maximum flow problem**. This observation is of independent interest because the maximum flow problem is a very intensively studied optimization problem, and thus everything that is known about the maximum flow problem can be brought to bear in analyzing the persuasion problem. Two interesting consequences of this relationship will be explored in Section 6.3:

1. Under normality, there always exist optimal persuasion rules that treat *ex ante* equivalent persuasion rules equivalently.

2. Comparative statics for the maximum flow problem can be adapted to the persuasion problem to show how the difficulty of persuading the listener changes as the probabilities of the states in A and R change.

In order to proceed it is necessary to explain the maximum flow problem in more detail, which is what I do next.

A **network** is a tuple (V, E, c, s, t), where (V, E) is a directed graph consisting of a set V of **vertices** and a set $E \subseteq V \times V$ of **directed edges**, $c : E \to \mathbb{R}_+ \cup \{\infty\}$ gives the capacity of each edge, $s \in V$ is the **source** while $t \in V$ is the **sink**. For any $v \in V$, define $\delta^+(v) := \{(u, w) \in E : u = v\}$ and $\delta^-(v) := \{(u, w) \in E : w = v\}$. A **flow** is a function $\varphi : E \to \mathbb{R}_+ \cup \{\infty\}$ satisfying $\varphi(e) \le c(e), \forall e \in E$ (the **capacity constraints**) and

$\sum_{e\in\delta^+(v)} \varphi(e) = \sum_{e\in\delta^-(v)} \varphi(e), \forall v \in V \setminus \{s,t\}$ (the **flow conservation constraints**). The capacity constraints say that the flow along any edge cannot exceed that edge's capacity, and the flow conservation constraints say that the flow into any vertex (other than the source or sink) is equal to the flow out of that vertex. The objective is to choose a flow to maximize the net flow out of the source:

$$\text{value}(\varphi) := \sum_{e\in\delta^+(s)} \varphi(e) - \sum_{e\in\delta^-(s)} \varphi(e) \tag{7}$$

The next step is to convert the persuasion problem into a maximum flow problem.[19] Start with a normal persuasion problem (X, M, σ, A, p), and construct the network $N = (V, E, c, s, t)$ such that $V = X \cup \{s, t\}$,

$$E = \underbrace{\{(s,x) : x \in A\}}_{E_1} \cup \underbrace{\{(x,y) : x \in A, y \in R, \sigma(x) \subseteq \sigma(y)\}}_{E_2} \cup \underbrace{\{(y,t) : y \in R\}}_{E_3}$$

$c(s,x) = p_x$ for all $(s,x) \in E_1$, $c(y,t) = p_y$ for all $(y,t) \in E_3$, and $c(e) = \infty$ for all $e \in E_2$. Observe that for normal message structures, $\sigma(x) \subseteq \sigma(y) \Leftrightarrow m_x \in \sigma(y)$, the edge set E_2 can be formed by checking for each $x \in A$ and $y \in R$, whether $m_x \in \sigma(y)$ (Theorem 4 tells us that the maximal message m_x at state x is easy to find). This makes it clear the conversion from the persuasion problem to the maximum flow problem can be done in polynomial time.

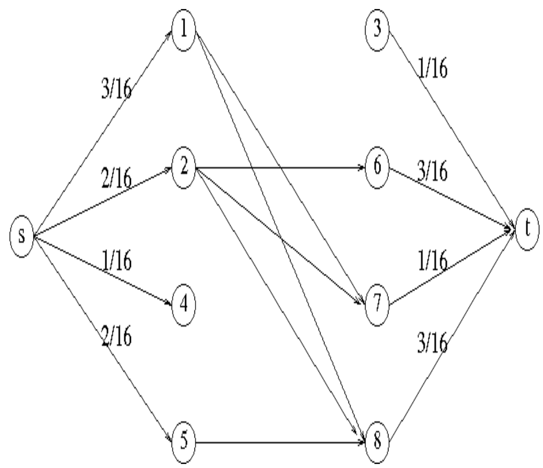
Example 5 I will now demonstrate this construction of the maximum flow problem corresponding to a persuasion problem in an example. Consider the persuasion problem given by the following table:

State	*A/R*	*Prob*	*Messages*
1	A	*3/16*	m_1, m_2, m_3
2	A	*2/16*	m_2, m_3
3	R	*1/16*	m_3
4	A	*1/16*	m_1, m_2, m_3, m_4, m_5
5	A	*2/16*	m_3, m_5
6	R	*3/16*	m_2, m_3, m_6
7	R	*1/16*	m_1, m_2, m_3, m_7
8	R	*3/16*	m_1, m_2, m_3, m_5, m_8

The first column represents the set of states $X = \{1, \ldots, 8\}$. The second column shows whether each state belongs to A or R. The third column gives the probability of each state,

[19] This construction is similar to the construction used by Picard (1976) to convert an instance of the maximal closure problem into an instance of the maximum flow problem. See also Hochbaum (2001) and Hochbaum (2004). The contribution of the current paper is to apply this to the persuasion problem, and in particular, show how this construction can be used to find an optimal persuasion rule and an *equilibrium* of the persuasion game.

and the last column gives the set of messages available at the state. The message structure is normal, and at each state $x \in \{1, \ldots, 8\}$, the maximal message is m_x. For instance m_1 is the maximal message at state 1. Message m_1 is available at state 4, and indeed all messages available at state 1 are available at state 4, as required by normality. We can check the analogous facts at each state. Another way to verify normality is given immediately after Theorem 4 at the end of Section 6. We now form the network corresponding to this persuasion problem:



The states in A and R correspond to vertices in the left and right columns respectively. The graph also contains a source s and a sink t. An edge with capacity p_x points from s to each state $x \in A$; an edge with capacity p_y points from each state $y \in R$ to t. For all $x \in A$ and $y \in R$, if x can mimic y (i.e., $\sigma(x) \subseteq \sigma(y)$, or equivalently $m_x \in \sigma(y)$) there is an edge with infinite capacity pointing from x to y. The infinite capacities are not labeled to avoid cluttering the diagram. $\square$

For any edge $e = (v, w)$, define $\overleftarrow{e} := (w, v)$ to be the edge pointing in the opposite direction. A network $N = (V, E, c, s, t)$ is **asymmetric** if for all edges e, $e \in E \Rightarrow \overleftarrow{e} \notin E$. Notice that any network corresponding to a normal persuasion problem is asymmetric. When discussing algorithms to solve the maximum flow problem, we restrict attention to asymmetric networks for ease of exposition, although such algorithms work for all networks with very minor modifications. For any asymmetric network $N = (V, E, c, s, t)$ and any feasible flow φ in that network define the **residual network** to be $N^\varphi = (V, E^\varphi, c^\varphi, s, t)$, where:

$$E^\varphi := \underbrace{\{e : e \in E, \varphi(v, w) < c(v, w)\}}_{E_1^\varphi} \cup \underbrace{\{e : \overleftarrow{e} \in E, \varphi(\overleftarrow{e}) > 0\}}_{E_2^\varphi}$$

$$c^\varphi(e) := \begin{cases} c(e) - \varphi(e), & \text{if } e \in E_1^\varphi\ ; \\ \varphi(\overleftarrow{e}), & \text{if } e \in E_2^\varphi. \end{cases}$$

Asymmetry of N implies that E_1^φ and E_2^φ are disjoint, and hence c^φ is well defined. Obviously, N^φ may not be asymmetric even though N is asymmetric. We refer to $c^\varphi(e)$ as the **residual capacity** of e. The basic idea behind the residual network is that any edge $e \in E$ has residual capacity equal to the capacity remaining on e, and for each edge $e \in E$, with positive flow $\varphi(e) > 0$ along e, the residual network contains a reverse edge $\overleftarrow{e}$ with residual capacity equal to the flow $\varphi(e)$; pushing additional flow along this reverse edge would correspond to undoing flow along the original edge. A **path** P in a graph (V, E) is a sequence of *distinct* vertices $(v_1, \dots, v_n)$ such that for all $i = 1, \dots, n-1$, $(v_i, v_{i+1}) \in E$. For any path $P = (v_1, \dots, v_n)$, and edge $e \in E$, $e \in P$ means that $e = (v_i, v_{i+1})$ for some $i = 1, \dots, n-1$. An **s-t-path** is a path in (V, E) from s to t; in other words the first vertex in the path is s and the last is t. For any feasible flow φ define:

$$V^\varphi := \{v \in V : \text{ there is a path from } s \text{ to } v \text{ in } (V, E^\varphi).\}^{20}$$

It is possible to show that a feasible flow φ is a maximum flow if and only if $t \notin V^\varphi$, or in other words, the sink is not reachable from the source in the residual network induced by φ. Let φ be a flow in an asymmetric network $N = (V, E, c, s, t)$, and let P be a path in (V, E^φ). Let γ be a real number. We say that φ is **augmented along** P **by** γ when we redefine $\varphi(e) := \varphi(e) + \gamma$ if $e \in P$ and $e \in E$, and redefine $\varphi(e) := \varphi(e) - \gamma$ if $e \in P$ and $\overleftarrow{e} \in E$. (Since N is asymmetric if cannot be the case that both e and $\overleftarrow{e}$ belong to E). $\varphi(e)$ is unchanged for all edges not in P. When we augment a flow along a path by γ, we increase the flow along edges in E that belong to the path by γ, and decrease the flow along edges in E whose reverse edges belong to the path by γ.

The following algorithm, known as the Ford-Fulkerson algorithm (Ford and Fulkerson 1957), finds a maximum flow in an asymmetric network.[21]

1. Set $\varphi(e) := 0$ for all $e \in E$.
2. Find an s-t-path P in N^φ. If none exists, then stop.
3. Let $\gamma := \min\{c^\varphi(e) : e \in P\}$. Augment φ along P by γ. Go to step 2.

On general graphs, the Ford-Fulkerson algorithm may require an exponential number of steps,[22] However, Edmonds and Karp (1972) solved these problems by showing that if at step 2, a shortest s-t-path is always chosen, then φ must be augmented at most $|V| \cdot |E|/2$ times and the algorithm runs in polynomial time.[23] In the case of graphs derived from

[20]Observe that (s) is path starting and ending at s. So $s \in V^\varphi$.

[21]By a slight modification of the above definitions, this algorithm will also find a maximum flow on an arbitrary–possibly not asymmetric–network. Note also that the algorithm was designed for networks in which all edge capacities are finite, whereas we have assumed that some edge capacities are infinite; however the problem would not be changed if all infinite capacities were replaced by sufficiently large finite capacities, and therefore, the algorithm is valid for the problem studied here.

[22]When edge capacities are *irrational*, it may not terminate at all.

[23]See Chapter 8 of Korte and Vygen (2006) for further details about facts in this paragraph.

normal persuasion problems, even without the solution of Edmonds and Karp (1972), the Ford-Fulkerson algorithm is guaranteed to run in polynomial time. This follows from the following observation.

Observation 1 *In graphs corresponding to a normal persuasion problem, the flow φ will be augmented at most $|X|$ times in the Ford-Fulkerson algorithm.*

The reason is that the edges in E that become saturated at Step 3 must belong to $E_1 \cup E_3$. (An edge becomes **saturated** when the flow along that edge is equal to its capacity). Moreover, one can show by induction that the flow along edges in $E_1 \cup E_3$ increases monotonically during the course of the algorithm. (This is not true of edges in E_2 whose flow may be reduced). Observation 1 now follows from the fact that $|E_1 \cup E_3| = |X|$.

The next theorem shows that the Ford-Fulkerson algorithm–or for that matter, any algorithm for the maximum flow problem–can be used to find: (1) an optimal persuasion rule, (2) a speaker strategy in a credible implementation of the optimal rule, (3) and the probability of error induced by the optimal rule.

Theorem 7 *Fix a normal persuasion problem satisfying $m_x = m_y \Rightarrow x = y$ for all $x, y \in A$.*[24] *Let φ be a maximum flow in the corresponding maximum flow problem. Then:*

$$f(m) := \begin{cases} 1, & \text{if } m = m_x \text{ for some } x \in A \cap V^{\varphi}; \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

is an optimal persuasion rule. A speaker strategy in the credible implementation is:

$$\zeta(x, m) := \begin{cases} 1, & \text{if } x \in A, m = m_x. \\ \frac{\varphi(y,x)}{\varphi(x,t)}, & \text{if } x \in R, m = m_y \in \sigma(x) \text{ with } y \in A; \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

except at $x \in R$ with $\varphi(x, t) = 0$, where $\zeta(x, m)$ can be chosen arbitrarily. The error probability is given by value(φ).

Proof. In Appendix. □

This theorem amounts to a constructive proof of the credibility result for normal problems. One important qualitative conclusion is that in every normal persuasion problem there is an optimal rule and a credible implementation of that rule such that at all states in A, the speaker tells the whole truth, and in states in R, the speaker randomizes (possibly

[24]This last assumption is innocuous because if it fails, we could always combine x and y into a single state z with $p_z = p_x + p_y$ and $\sigma(z) = \sigma(x) = \sigma(y)$ (noting that the last equality follows from normality and $m_x = m_y$), arriving at an equivalent persuasion problem.

degenerately) over lies (unless he cannot lie).[25] By the definition of a credible implementation, the theorem only states that (ζ, f) is a Bayesian Nash equilibrium. (ζ, f) is also a sequential equilibrium unless there is some message m which can only be sent in states in A and such that $f(m) = 0$. In this case simply reset $f(m) = 1$ for any such message and then (ζ, f) is a sequential equilibrium as well. Since f was optimal before this change, it must still be optimal after this change, and it follows that there cannot be any state $x \in A$ at which the speaker prefers to play any message other than m_x even after the change. For more information on this issue, the reader is referred to Theorem 2 and its proof.

Example 5 continued Let us apply the Ford-Fulkerson algorithm to the network in Example 5:

1. Find the path from s to 2 to 8 to t in the original graph. The minimum capacity edge is $(s, 2)$, which has capacity 2/16. So push 2/16 units of flow along this path.

2. Find the path from s to 1 to 8 to t in the residual graph. The minimum capacity edge in the residual graph is $(8, t)$, which has residual capacity 1/16 (since 2/16 was added to this edge at the previous step). So push 1/16 units of flow along this path.

3. Find the path from s to 5 to 8 to 2 to 6 to t. Notice that in moving from 8 to 2, we are moving backwards along the edge $(2, 8)$, which received 2/16 units of flow in the first step. Recall that we can move backwards along an edge when there is already flow on that edge, and the residual capacity on this backwards edge $(8, 2)$ is equal to the flow along $(2, 8)$. Pushing flow along $(8, 2)$ corresponds to undoing the flow on $(2, 8)$. In fact, the two minimum capacity edges in the residual graph along this path are $(s, 5)$ and $(2, 8)$ both of which have a capacity of 2/16. Therefore, we push a flow of 2/16 along this path.

4. Find the path from s to 1 to 7 to t in the residual graph. The minimum capacity edge in this path is $(7, t)$. It has capacity 1/16. Push one 1/16 units of flow along this path.

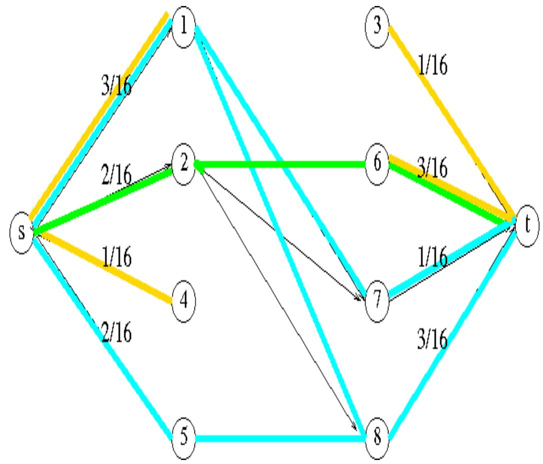
5. The sink is no longer reachable from the source in the residual graph, so the algorithm terminates with the flow:

$(s,1)$	$(s,2)$	$(s,4)$	$(s,5)$	$(1,7)$	$(1,8)$	$(2,6)$	$(2,7)$	$(2,8)$	$(5,8)$	$(3,t)$	$(6,t)$	$(7,t)$	$(8,t)$
2/16	2/16	0	2/16	1/16	1/16	2/16	0	0	2/16	0	2/16	1/16	3/16

The top column gives the edge, and the entry in the bottom column below an edge gives the flow along that edge. For example the flow $\varphi(s, 1)$ along $(s, 1)$ is equal to 2/16. Notice that several times the algorithm found a certain path in the original or residual graph. The exact details about how this was done were omitted, but could easily have been filled in.

[25] Notice that the term "lie" here is used more broadly than in 2 in the introduction to include any message in which the speaker does not tell the whole truth.

The exact course of the algorithm and the number of steps depends on this. For example, we could have started with the path from s to 1 to 8 to t. This would have led to a different maximum flow. The flow which we derived above is illustrated in the following figure:



Recall that all edges in the original network point from left to right, that is in the direction away from s and towards t. The blue and green lines depict flow along an edge. The flow is blue if it passes through vertices corresponding to states that are accepted by the corresponding persuasion rule, and it is green if it passes through vertices that are rejected by the corresponding persuasion rule. The yellow lines depict residual capacity on the finite capacity edges. (Yellow lines have been omitted from the infinite capacity edges in the middle to avoid cluttering the graph). Even without having observed the exact steps of the algorithm we can verify that this is a maximum flow because the the sink is not reachable from the source in the residual graph induced by the flow. For example, there is residual capacity on $(s, 1)$ and on $(1, 8)$ (since the latter has infinite capacity), so it is possible to reach 8 from s in the residual graph, but then it is not possible to reach t from 8, because there is no residual capacity on $(8, t)$, so t is not reachable from s in this way. Similarly, one can verify that there is no path from s to t in the residual graph. Indeed, the algorithm terminated for precisely this reason.

Since the above is a maximum flow φ, Theorem 7 implies that an optimal persuasion rule accepts messages in $\{m_1, m_4, m_5\}$ and rejects all others, because vertices $1, 4$, and 5 are precisely the vertices in A which are reachable from the source s in the residual graph. It is obvious that 1 and 4 are reachable from s because there is residual capacity on the edges $(s, 1)$ and $(s, 4)$. But there is no residual capacity on the edge $(s, 5)$, so how can we reach 5 in the residual graph? We travel from s to 1 to 8 to 5. Recall that the edge $(1, 8)$ has infinite capacity; moreover, we travel on $(8, 5)$, which is the reverse edge of $(5, 8)$, and so $(8, 5)$ has capacity equal to the flow along $(5, 8)$ which is $2/16$. On the other hand, notice that 2 is not reachable from s in the residual graph, so the message m_2 is not accepted by the optimal rule corresponding to φ.

Theorem 7 also tells us how to use the flow φ to identify the speaker strategy in a credible implementation of the optimal rule f which accepts exactly messages m_1, m_4, and m_5. In particular, at all states x in A, the speaker sends m_x. Consider next the states in R. Since $\varphi(3, t) = 0$, the theorem tells us that the speaker's strategy can be chosen arbitrarily at this state, and indeed the only option is for the speaker to send message m_3 since this is the only message available at x_3. Theorem 7 also tells us that at state 6, the speaker sends message m_2 with probability 1, at state 7, the speaker sends message m_1, with probability 1, and at state 8, the speaker sends messages m_1 and m_5 with probabilities proportional to the flow along the edges $(1, 8)$ and $(5, 8)$ respectively. In particular, this means that at state 8, the speaker sends m_1 with probability 1/3 and m_5 with probability 2/3.

We can easily verify that given this speaker best best reply, the optimal rule f corresponding to φ is also a best reply to the speaker's strategy. For example, if the listener receives message m_5, he knows that the state is either 5, which is in A or 8, which is in R. State 5 occurs with probability 2/16, and state 8 with a higher probability 3/16, but at state 5 the speaker sends m_5 with probability 1, whereas at state 8, the speaker only sends m_5 with probability 2/3, so conditional on m_5 being sent, the probability that the state is 5 is equal to the probability that the state is 8; both probabilities are 1/2. So the listener is exactly indifferent between accepting and rejecting m_5, and is therefore willing to accept as is called for by rule f. As another example, if m_1 is received, the listener can conclude that the state is either 1, which is in A or 7 or 8, both of which are in R. State 1 occurs with probability 3/16 and in state 1, the speaker sends m_1 with probability 1. State 7 occurs with probability 1/16 and in state 7, the speaker also sends m_1 with probability 1. State 8 occurs with probability 3/16, but in state 8, the speaker only sends m_1 with probability 1/3, so the probability that the state is 1 is 3/5, which is greater than 1/2, so the listener strictly prefers to accept message m_1, as required by f.

Finally Theorem 7 tells us that the error probability at the optimal rule f is equal to 6/16, the value of the flow, which is equal to the amount of flow emanating from the source, or equivalently, the amount of flow pouring into the sink. The reason for this is as follows: the maximum flow matches up probability mass from states the listener would like to accept with probability mass with states the listener would like to reject that can mimic the states he would like to accept. The listener must make a mistake on one side or another. This also helps us to understand why it is optimal for the speaker to accept message m_1, for example. Even after we match up 2/16 of the probability mass with probability mass of states in R that can mimic state 1, and have matched up the remaining probability mass of other states in R which can mimic 1, with probability mass of other states in A that they can mimic, there is still capacity 1/16 left on edge $(s, 1)$, or equivalently probability mass 1/16 left unmatched of the full probability mass of state 1, and this represents the net benefit of accepting message m_1.

6.3 Qualitative Properties of Persuasion

In Section 6.2, a relationship between the persuasion problem and a classical optimization known as the maximum flow problem was used to show that any algorithm solving the maximum flow problem can also be used to solve the persuasion problem under normality. In this section, I show that this same relationship can be used to derive various qualitative properties of the solution. The key property of the maximum flow problem, which leads to these qualitative results is that it is an instance of supermodular optimization (Topkis 1998).

6.3.1 Symmetry

Example 3 shows that messages which have equivalent evidentiary content may be treated differently by optimal persuasion rules. In that example, evidence that two experts of the same gender support the speaker's position is persuasive, but the opinions of two experts of different genders is not. Yet gender is irrelevant. A similar example resembles Fishman and Hagerty (1990): a speaker attempts to persuade a listener that the speaker is talented in some task. The listener knows that the speaker will take five attempts at the task regardless of his performance. On each attempt, the speaker will be either successful or unsuccessful. Conditional on his talent, his performance on the tasks is i.i.d.. The speaker will only report the outcome of one task, but cannot lie. Intuitively, it still may be more persuasive for the speaker to say "I was successful on the first attempt", than "I was successful on the fourth attempt," even though in a setting in which the speaker randomly chooses which outcome to report neither should be more persuasive. This may be because it is common knowledge that the listener expects the speaker to report the first task on which he was successful. If the speaker reports the first successful task because it is more persuasive, the listener would acquire more information than if the listener were to respond to any claim of the form "I was successful on attempt n" in the same way and the speaker randomly chose a success to report (whenever he was successful on some attempt). Both of the examples above contain a constraint on the amount of evidence which the speaker reports. This turns out to be essential to the asymmetric treatment of identical evidence. I will now formalize this observation.

Definition 2 *A pair of bijections (π, ξ) where $\pi : X \to X$ and $\xi : M \to M$ is a* ***symmetry*** *if for all x: (i) $\sigma(\pi(x)) = \{\xi(m) : m \in \sigma(x)\}$, (ii) $x \in A \Leftrightarrow \pi(x) \in A$, and (iii) $p_{\pi(x)} = p_x$. A persuasion rule f is* ***symmetric*** *if for every symmetry (π, ξ), $f = f \circ \xi$.*

A symmetry is a pair of functions, one from states to states, and the other from messages to messages, which preserve all relevant properties of the model. A persuasion rule is symmetric if it treats any pair of messages which cannot be distinguished without labels–or in other words, in terms of their intrinsic properties–in the same way. The two examples described above involving (i) gender (Example 3) and (ii) reporting only one task showed how optimal persuasion rules may be asymmetric.

Theorem 8 *Every normal persuasion problem admits a symmetric optimal persuasion rule.*

Proof. See Appendix. □

Theorem 8 follows from the supermodularity of the persuasion problem under normality. Observe that Theorem 8 allows for the possibility that some optimal persuasion rule is asymmetric, but under normality this phenomenon is inessential in the sense that whenever there is an asymmetric optimal rule, there is also a symmetric optimal rule. This establishes that time constraints, leading to a failure of normality, are critical for asymmetric treatment of equivalent evidence in an essential manner.

I will now present a class of non-normal persuasion problems, which I will refer to as the class of **typical time constrained** persuasion problems, in which all optimal rules are asymmetric whenever the problem is not trivial. This class includes the two examples discussed in the beginning of this section as special cases. In particular, suppose that $X = \{0,1\}^n$. For any $J \subseteq \{1,\ldots,n\}$, and $x = (x_1,\ldots,x_n) \in X = \{0,1\}^n$, define $x_J = \{(x_j, j) : j \in J\}$, and suppose that $\sigma(x) := \{x_J : |J| \leq h\}$. This means that at any state, the speaker can reveal (at most) h components of the vector he observes. If the speaker reveals a component, he reveals both its value (in $\{0,1\}$) and its index (in $\{1,\ldots,n\}$). Assume that $1 \leq h < n$, so at x, the speaker can show at least 1 but not all of the components. Assume there exists some function $g : X \to [0,1]$ such that for all x, $p_x = g(\sum_i x_i)$, so the probability of a string depends only on the number of 1's and 0's in the string. For example, this would be true if the components of x are n i.i.d. draws from a binary random variable, or if first some random variable θ is realized, and then x is the realization of n draws which are i.i.d. conditional on θ. Assume moreover that for some ℓ with $h < \ell \leq n$, $A = \{x \in X : \sum_i x_i \geq \ell\}$. In other words, there is some critical number ℓ such that if the speaker has at least ℓ components equal to 1, then the listener would like to accept the speaker's request, and otherwise, the listener would like to reject it. ℓ is sufficiently large, that it is always impossible for the speaker to prove that he has ℓ components equal to 1.

Theorem 9 *In typical time constrained persuasion problems, if it is not optimal to reject every message, then every optimal persuasion rule is asymmetric.*

Proof. In Appendix.

Observe that it is critical for the results in this section that the listener knows which action the speaker prefers. Imagine, in contrast a situation of "cheap talk" in which every message is available in every state of the world. Suppose there are two actions, a and b, and that the speaker and listener's interests are perfectly aligned in every state. Also, the optimal action depends on the state which is privately known to the speaker. Any optimal persuasion rule would employ at least two messages, which would be treated differently. For example, if a were optimal, the speaker would report one message, say m^a, and if b were optimal, the speaker would report another message, say m^b, and the listener would respond to m^a with a and to m^b with b. The message structure is trivially normal, and every optimal persuasion rule is asymmetric, because all messages are ex ante the same, so a symmetric persuasion rule would have to treat them in the same way. A little reflection shows that asymmetric treatment of equivalent messages is essential to all linguistic communication. For example, one word is used to refer to a house in English, and a different word is used in French. These two words are ex ante equivalent; before a language develops, it does not matter which one will be used to refer to a house. The systematically different treatment of arbitrary signs is what allows people to communicate. Why then is the symmetry result in Theorem 8 relevant? The answer is that we are dealing with a situation in which the speaker's preferences are common knowledge. If there were only cheap talk, meaningful communication would be impossible, because at every state the speaker would use the most persuasive message, which would not allow the listener to infer any information about the state. So what allows for communication in this case is the presence of evidence whose availability depends on the state. The question then becomes, in such a setting is there any reason for the listener to treat equivalent pieces of evidence differently, and the above theorems show that the answer is: with time constraints, yes, but without time constraints, no.

6.3.2 Comparative Statics

Persuasion rule f is **more difficult** than persuasion rule f' if for all $x \in X$, $\alpha(f,x) \leq \alpha(f',x)$. A more difficult persuasion rule is one such that in each state the speaker is less persuasive, and hence worse off. A persuasion rule f is a **most difficult optimal rule** if f is optimal and for all optimal f', f is (weakly) more difficult than f'. If f and f' are both most difficult optimal rules, then $\alpha(f) = \alpha(f')$. **Least difficult optimal rules** are defined similarly. Rules that are either most or least difficult optimal rules are **extreme**.

Theorem 10 *In a normal persuasion problem, there exist most and least difficult optimal rules. The extreme optimal rules are deterministic. Shifting probability mass from a state $x \in A$ to a state $y \in R$ will make the extreme optimal rules (weakly) more difficult. Moving a state x from A to R will have the same effect. Without normality, most and least difficult optimal rules may not exist; even when there is a unique optimal persuasion rule f, shifting probability mass from a state $x \in A$ to a state $y \in R$ or moving a state x from A to R may lead to a persuasion rule f' which is neither more nor less difficult than f.*

Proof. In Appendix. $\square$

It may be tempting to conclude from Theorem 10 that in a normal persuasion problem, shifting probability mass from a state $x \in A$ to a state $y \in R$ will result in a higher probability of acceptance. However, the following example shows that this is false.

Example 6 Suppose that $X = \{1,2,3,4\}$, $A = \{1,2\}$, $M = \{o,e\}$, $\sigma(1) = \sigma(3) = \{o\}, \sigma(2) = \sigma(4) = \{e\}$. (X, M, σ) is a normal message structure. Assume that:

$$\begin{array}{llll} p_1 = 1/3 & p_2 = 2/9 & p_3 = 1/9 & p_4 = 1/3 \\ p'_1 = 1/3 & p'_2 = 1/9 & p'_3 = 2/9 & p'_4 = 1/3 \end{array}$$

Then in moving from $(p_x)_{x\in X}$ to $(p'_x)_{x\in X}$ probability mass is shifted from state 2 in A to state 3 in R. However in both cases the unique optimal persuasion rule accepts o and rejects e, and the probability of acceptance moves up from 4/9 to 5/9.

This example should be contrasted with the following:

Theorem 11 *In a normal persuasion problem, moving a state x from A to R reduces the probability of acceptance at the extreme optimal rules.*

Proof. By Theorem 10, the extreme optimal rules are deterministic, and moving x from A to R cases these rules to become (weakly) more difficult, that is accept the speaker at fewer states. The conclusion follows from the fact that the probabilities of states have not changed. □

Without normality, it is possible that in a situation with a unique optimal rule, in which moving a state x from A to R may increase the acceptance probability.[26]

Fixing all aspects of the persuasion problem except $p = (p_x)_{x\in X}$, let $\mu(p)$ be the error probability of error at the optimal rule as a function of p. Moreover, let $(p_{\overline{x}}+\epsilon, p_{\underline{x}}-\epsilon, p_{-\{\underline{x},\overline{x}\}})$ be the probability distribution on X that results by starting from p and shifting ϵ probability mass from state $\underline{x}$ to state $\overline{x}$.

Theorem 12 *Fix a normal message structure (X, M, σ) and a set A. Choose $\overline{x} \in A, \underline{x} \in R$. Let $p = (p_x)_{x\in X}$ and $p' = (p'_x)_{x\in X}$ be such that $p_x \leq p'_x$ for all $x \in A$ and $p'_x \leq p_x$ for all $x \in R$. Moreover, let $\epsilon \in [0, \min\{p_{\underline{x}}, p'_{\underline{x}}\}]$. Then:*

$$\mu(p_{\overline{x}} + \epsilon, p_{\underline{x}} - \epsilon, p_{-\{\underline{x},\overline{x}\}}) - \mu(p) \geq \mu(p'_{\overline{x}} + \epsilon, p'_{\underline{x}} - \epsilon, p'_{-\{\underline{x},\overline{x}\}}) - \mu(p') \tag{10}$$

This theorem is similar to Theorem 3.7.3 in Topkis (1998), and so the proof is omitted. The theorem states that the shifting a fixed probability mass from a state in R to a state in A reduces the probability of error by more (or increases it by less), after some probability

[26]This can be seen by considering the following example: $X = \{1,2,3,4\}, M = \{m_1, m_2, m_3\}, \sigma(1) = \{m_1, m_2\}, \sigma(2) = \{m_3\}, \sigma(3) = \{m_1\}, \sigma(4) = \{m_2, m_3\}, A = \{1,2,3\}, p_1 = 3/8, p_2 = 1/8, p_3 = p_4 = 1/4$. Consider what happens when 3 is moved from A to R.

mass has already been shifted from states in R to states in A. Counter-examples to the theorem can be constructed without normality.

7 What Must We Know to Make The Persuasion Problem Tractable?

This section addresses the question of why the persuasion problem is difficult without normality. The approach is to ask the following question: what additional information would we need in order for it to be easy to find an optimal persuasion rule? The answer is: a speaker best reply to the optimal rule. Notice that in normal persuasion problems, we always know a speaker best reply to an optimal rule *a priori*:

Theorem 13 *In a normal persuasion problem, there always exists an optimal persuasion rule such that at all states x it is a best reply for the speaker to send m_x.*

Proof. The optimal rule f^* constructed in Theorem 7 is such that for all $x \in A$ it is a best reply for the speaker to send m_x. Now alter f^* so that for all $x \in R$, if there exists $y \in A$ with $f(m_y) = 1$ and $m_y \in \sigma(x)$, then redefine $f^*(m_x) = 1$. f^* is still optimal under this alteration, and it is now a best reply for the speaker to send m_x at all states x. $\square$

The complexity of an finding an optimal persuasion rule has to do with the number of computations necessary to find that rule. However, suppose that we have a "black box" that will answer a certain type of yes-no question for us for free. In complexity theory, such a black box is called an **oracle**. Let $Y = \{x_1, x_2, \ldots, x_n\} \subseteq A$ and for $\{m_1, m_2, \ldots, m_n\}$ be any collection of messages such that $m_i \in \sigma(x_i)$. Then consider the following oracle:

Best Reply Oracle

Does there exist an optimal persuasion rule f^* such that x_1 sends m_1, x_2 sends $m_2, \ldots,$ and x_n sends m_n in some best reply to f^*?

Notice that the best-reply oracle only needs to answer questions about the speaker's strategy in states in A. A problem which is hard may become easy if we can appeal to the oracle. The oracle may greatly reduce the number of steps in the computation, since we can simply ask the oracle to instantly perform certain hard computations. For a formal development of oracles, see Papadimitriou (1994).

Theorem 14 *With the best-reply oracle, there exists a polynomial time algorithm that finds an optimal persuasion rule and a credible implementation of the optimal rule.*

The proof and further details are given in the appendix. The outline of the proof is as follows. Start with an arbitrary persuasion problem $Q = (X, M, \sigma, A, p)$. Use the oracle to find a speaker best reply ζ to some optimal rule for states x in A. Now consider a related persuasion problem $Q^* = (X, M, \sigma^*, A, p)$ in which in states in A, the speaker has only one choice: she must send the message that she sends according to ζ. In states in R, the set of available messages is unchanged. In all other respects, Q and Q^* are the same. Although Q^* is not necessarily normal, it satisfies a weaker condition **normality on A**:

$$\forall x \in A, \exists m_x \in \sigma(x), \forall y \in X, m_x \in \sigma(y) \Rightarrow \sigma(x) \subseteq \sigma(y)$$

This condition is trivially satisfied in Q^* because for all $x \in A$, $\sigma^*(x)$ is a singleton. Normality on A turns out to be sufficient for translating the persuasion problem into a maximum flow problem. One can then use the maximum flow algorithms to find an optimal persuasion rule f and credible implementation (ζ, f) in Q^* as in Theorem 7. Finally, one shows that f and (ζ, f) are also an optimal persuasion rule and a credible implementation in the original problem Q.

Theorem 14, when combined with Theorem 13, helps us to understand why it is difficult to solve the persuasion problem without normality. In particular, without normality, we do not know *a priori* what messages the speaker will send in a best reply to the optimal rule. We must simultaneously discover what reporting strategy ζ the speaker will ultimately use and an optimal persuasion rule which makes ζ a best reply. In contrast, Theorem 13 shows us that in a normal persuasion problem we can assume that the speaker tells the whole truth, and we must search for an optimal persuasion rule among those which make truth-telling a best reply. This is analogous to the revelation principle in ordinary mechanism design: we search for the optimal mechanism among *incentive-compatible* mechanisms, where incentive compatibility means that truth telling is a best reply.[27] However, in non-normal persuasion problems, we cannot make any *a priori* assumption about what the speaker's best reply to the optimal rule will be. This leads to a hard optimization problem, as attested to by Theorem 3. If we had an oracle that told us the speaker's best reply to the optimal rule, then Theorem 14 shows us that the problem would become tractable again.

8 Cheap Talk and Dynamic Communication

It is well known that the analysis of the properties of Nash equilibrium is more computationally intractable than the analysis of correlated equilibrium (Gilboa and Zemel 1989). A correlated equilibrium of a normal form game is a Nash equilibrium of an augmented game which includes a disinterested mediator and unlimited communication possibilities between the mediator and the players of the original game. The analogue of a correlated equilibrium

[27] Several papers have shown that normality and other similar properties are critical for versions of the revelation principle in mechanism design environments with evidence. (Green and Laffont 1977, Forges and Koessler 2005, Bull and Watson 2007, Deneckere and Severenov 2008)

in the case of a Bayesian game is a communication equilibrium. The reason that correlated equilibrium is more computationally tractable is that the set of correlated equilibria of a game–unlike the set of Nash equilibria–is defined by a set of linear inequalities. Finding the best correlated equilibrium for a player in a normal form game amounts to solving a linear program, whereas finding the best Nash equilibrium for some player does not. Indeed, Gilboa and Zemel (1989) showed that it is NP-hard to find a Nash equilibrium in an arbitrary two-player normal form game that maximizes the sum of utilities of the players, and Conitzer and Sandholm (2008) showed that it is NP-hard to find a Nash equilibrium that maximizes the utility of one of the players.

The credibility result implies that finding an optimal persuasion rule amounts to finding the listener's strategy in the best Nash equilibrium (also best sequential equilibrium–see Theorem 2) for the listener in the game without commitment. The discussion of the previous paragraph suggests that if the game is augmented with additional communication possibilities, then it might be possible to achieve a computational simplification in finding the optimal persuasion rule. The persuasion problem studied in this paper assumes a very simple interaction: the speaker sends the listener a message, and then the listener takes an action. As an alternative, consider the following:

Dynamic Communication Protocol

1. The true state x is realized, and is known only to the speaker.
2. The speaker makes a cheap talk claim that the the state is $\hat{x}$.
3. The listener requests that the speaker present some message $m \in \sigma(\hat{x})$; the listener chooses his request randomly according to some probability distribution over $\sigma(\hat{x})$.
4. The speaker presents some message $m' \in \sigma(x)$.
5. (a) If the speaker fails to present the requested message (i.e., $m' \neq m$), the listener rejects the speaker's request.
   (b) If the speaker presents the requested message (i.e., $m = m'$), the listener accepts the speaker's request with some probability $q \in [0, 1]$, depending on the entire sequence of communication.

Notice that the listener must make a decision at two places above: Step 3 and Step 5b.[28] A **dynamic persuasion rule**–to be defined more formally below–specifies these decisions. This is to be contrasted with the persuasion rules f defined in Section 3, which I have studied throughout this paper, which simply specified which messages to accept and which messages

[28] No choice is necessary at Step 5a; the listener always rejects in this case. One can show that the listener could not do better if he had the option to accept at this point.

to reject. I will henceforth refer to such persuasion rules f as **static persuasion rules**. Recall that in Section 3 a **persuasion problem** was defined as a tuple (X, M, σ, A, p). We can now discuss both the optimal static persuasion rule and the optimal dynamic persuasion rule in a given persuasion problem. In this new terminology, all previous theorems concern optimal static persuasion rules. When discussing optimal dynamic persuasion rules, I will assume–as in the case of optimal static persuasion rules–that the listener can commit to a dynamic persuasion rules, and that the speaker best responds to this persuasion rule. Notice that a speaker's strategy now consists of two parts. In Step 2, the speaker sends a cheap talk message which may be any $\hat{x} \in X$. In Step 4, the speaker must send a message $m \in \sigma(x)$, where x is the true state. Notice a sharp contrast between these two decisions. Whereas communication in Step 4 is restricted in a way that depends on the true state, communication in Step 2 is unrestricted. Since the focus of this section is the complexity of finding optimal persuasion rules when the game is augmented by dynamic cheap talk, I will not address the issue of credibility for dynamic persuasion rules.[29]

Before proceeding with the discussion, it is important to discuss the interpretation of the dynamic communication protocol. This paper studies persuasion with *limited* communication. Indeed these communication limitations are what make the game interesting; since the listener knows the speaker's preferences, without communication limitations, no communication is possible. Augmentation with cheap talk may computationally simplify the persuasion problem precisely when the message structure is not normal; under normality, the persuasion problem is already tractable without cheap talk (Theorem 6). The interpretation of a failure of normality is that there are time constraints–or other similar constraints, such as cognitive or attentional constraints imposed by either the speaker or listener–that limit the speaker's ability to present *all* of her information. Since time is limited, the speaker must be selective in deciding which evidence to present. However, in the dynamic communication protocol, the speaker makes a cheap talk claim which provides *all* of her information prior to presenting some part of the evidence. There is no obvious reason why it should require less time or fewer cognitive resources to make a "cheap talk" claim than to present evidence.[30] Indeed, if time is limited, then more time spent talking about the evidence may leave less time to present the evidence. Note, moreover that as explained in the introduction, the messages governed by the message correspondence σ needn't be interpreted as hard evidence, but may have other interpretations as well; the important point is that there is some form of communication limitation. Evidence does appear to be the most natural interpretation in the context of the dynamic communication protocol. With these caveats in mind, I will proceed to discuss the dynamic communication protocol. When studying persuasion with limited communication, part of the explicit topic of interest is how

[29] In fact, one can prove a credibility result for dynamic persuasion rules.

[30] The term "cheap talk" may not really be appropriate in limited communication contexts; all communication may be costly.

different assumptions about communication and communication limitations–including the communication protocol–translate to different properties of persuasion. Therefore it is of interest to study the problem under different assumptions–in this case, the assumption that unlimited cheap talk can precede the presentation of evidence.

The dynamic communication protocol described above resembles communication protocols used in Bull and Watson (2007) and Glazer and Rubinstein (2004), which is a predecessor of Glazer and Rubinstein (2006).[31] It is possible to prove that (i) adding more

[31] Here I discuss the relationship between the above communication protocols and the communication protocols in Glazer and Rubinstein (2004) and Bull and Watson (2007). Glazer and Rubinstein (2004) assumed that X has a product structure so that $X = X_1 \times X_2 \times \cdots \times X_n$. In that paper, there was no presentation of evidence; rather the listener could verify some aspects of the speaker's claim. More precisely, first the true state $x = (x_1, \ldots, x_n)$ is realized, and is known only to the speaker. Then the listener makes a cheap talk claim that the state is $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_n)$. The listener then randomly selects one of the components i in $\{1, \ldots, n\}$ according to some probability distribution, and verifies that the speaker told the truth about the ith component (i.e., $\hat{x}_i = x_i$). If the speaker lied, the listener rejects his request. If the speaker told the truth, the listener accepts his request with some probability in $[0, 1]$. I now show that it is possible to replace verification by presentation of evidence, and therefore to view the model of Glazer and Rubinstein (2004) as presenting a special case of the dynamic communication protocol described above. In particular, construct a message correspondence such that for all $x = (x_1, \ldots, x_n) \in X$, $\sigma(x) := \{(i, x_i) : i = 1, \ldots, n\}$. In other words, the for any component, $i \in \{1, \ldots, n\}$, it is possible for the speaker to prove what the ith component of the state is, but the speaker may not provide evidence for the values of more than one component. The dynamic communication protocol described above with this message correspondence is essentially the same as the protocol in Glazer and Rubinstein (2004); instead of directly verifying that component i is as the speaker claims, the listener requests that the speaker prove that component i is as she claims. It is easy to see that the dynamic communication protocol described in this section is a *strict* generalization of the model in Glazer and Rubinstein (2004) as there are message correspondences σ which cannot be interpreted as presenting a single dimension of a multidimensional state. In contrast to the model in Glazer and Rubinstein (2006), optimal persuasion rules in Glazer and Rubinstein (2004) may inherently involve randomization on the listener's part. This is one of the consequences of dynamic communication.

The dynamic communication protocol also resembles the "special three-stage mechanism" of Bull and Watson (2007), which did not explicitly study persuasion, but rather studied general mechanism design problems with evidence. Bull and Watson (2007) constructed a framework for studying mechanism design problems with evidence. Theorem 6 of that paper showed that within the Bull and Watson (2007) framework, any dynamic outcome function (i.e., functions from type profiles to outcomes–such as *accept* and *reject*, for example) which is implementable by some mechanism is implementable by a special three stage mechanism in which all agents report truthfully in the cheap talk stage. The framework of Bull and Watson (2007) makes some implicit assumptions about the interpretation of the message correspondence σ which I do not commit to in this paper. It is important to note that because Bull and Watson (2007) focus on deterministic mechanisms, their theorem cannot be applied directly to the model analyzed in this section. In particular, when the listener can commit, we are trying to implement outcome functions in the one-player game in which the only player is the speaker. This one-player game is vacuously a game of complete information; when there is only one player it is trivial that all information is common knowledge. Suppose that we are trying to implement a deterministic outcome function in the persuasion game; it then follows from Theorem 6 of Bull and Watson (2007) that we can eliminate Step 3 in the dynamic communication protocol (in Bull and Watson (2007), this is the second stage). The listener does not request any specific evidence and simply responds to the evidence presented by the speaker. However, it is also easy to see by appealing to the fact that the speaker's preferences do not depend on the state that once Step 3 is eliminated, Step 2 can be eliminated as well. But then we are back to the model of Glazer and Rubinstein (2006)–or equivalently, the model that I present in Section 3. In that model, Glazer and Rubinstein (2006) have shown that randomization never helps the listener, but in Glazer and Rubinstein (2004) it was shown that randomization sometimes does make a difference, and as I have argued the model of Glazer and Rubinstein (2004) is a special case of the dynamic communication protocol that I present in this section. It follows that since we are interested in possibly random optimal persuasion rules, the conclusions of Bull and Watson (2007) on complicated sequences of communication prior to the presentation of evidence would not help the listener, (ii) given the listener commits to a rule, we may assume that the listener tells the truth in Step 2, (iii) it would not help the listener to be able to ask for messages which do not belong to $\sigma(\hat{x})$ in step 4, (iv) nor would it help the listener relax the assumption in Step 5a that the listener rejects the speaker's request if the speaker fails to present the requested evidence.[32]

I now describe dynamic persuasion rules more formally. A dynamic persuasion rule consists of a pair of functions (d, g), where $d : X \times M \to [0,1]$ and $g : X \times M \to [0,1]$. For all $x \in X, \sum_{m \in M} d(x,m) = 1$. $d(x,m)$ is the probability that the listener requests that the speaker shows him message m if the speaker claims that the state is x at step 2. We assume that for all $x \in X$ and $m \in M$, $d(x,m) > 0 \Rightarrow m \in \sigma(x)$, or in other words, the listener only asks for evidence which the speaker claims to have. $g(x,m)$ is the probability that the listener accepts the speaker's request conditional on the event that the speaker claimed that the state is x, the listener requested message $m \in \sigma(x)$ and the listener succeeded in presenting m. We adopt the convention that $d(x,m) = 0 \Rightarrow g(x,m) = 0$.[33] Note that $g(x,m)$ is only a function of of the speaker's claim about the state x and the listener's request to see evidence m, but g is not a function of the evidence that the speaker actually presents to the listener. This is because, if the speaker fails to present the requested evidence, it is assumed that the listener simply rejects his request, so $g(x,m)$ only gives the probability of acceptance if the speaker succeeds. In evaluating the error probability of a persuasion rule (d, g), we assume that the speaker best responds to (d, g).

Theorem 15 *Let* $(\beta^*(x,m) : x \in X, m \in \sigma(x))$ *solve*

$$\begin{aligned}
\min \quad & \sum_{x \in A} (1 - \sum_{m \in \sigma(x)} \beta(x,m)) p_x + \sum_{x \in R} \sum_{m \in \sigma(x)} \beta(x,m) p_x \\
\text{s.t.} \quad & \sum_{m \in \sigma(x)} \beta(x,m) \geq \sum_{m \in \sigma(x) \cap \sigma(y)} \beta(y,m), \quad \forall x, y \in X \qquad (11) \\
& \sum_{m \in \sigma(x)} \beta(x,m) \leq 1, \quad \forall x \in X \\
& \beta(x,m) \geq 0, \quad \forall x \in X, \forall m \in \sigma(x)
\end{aligned}$$

implementation of deterministic outcome functions do not apply directly.

Finally, note that all papers discussed here as well as the dynamic communication protocol can be embedded into the framework of Myerson (1986). However, again this may involve some implicit assumptions about the interpretation of the message correspondence σ.

[32] (i), (ii), and (iii) are similar in spirit to Theorem 6 of Bull and Watson (2007); however see footnote 31 for a discussion of why the results of that paper cannot be directly applied to the present model.

[33] In particular, since $d(x,m) > 0 \Rightarrow m \in \sigma(x)$, it follows that $m \notin \sigma(x) \Rightarrow g(x,m) = 0$.

Then the dynamic persuasion rule (d, g) given by:

$$\begin{aligned} d(x,m) &= \begin{cases} \frac{\beta^*(x,m)}{\sum_{m' \in \sigma(x)} \beta^*(x,m')}, & \text{if } \sum_{m' \in \sigma(x)} \beta^*(x,m') > 0; \\ 0, & \text{otherwise.} \end{cases} \\ g(x,m) &= \sum_{m' \in \sigma(x)} \beta^*(x,m') \end{aligned} \tag{12}$$

is optimal.[34]

Proof. In Appendix. □

Corollary 3 *There exists a polynomial-time algorithm for finding the optimal dynamic persuasion rule.*

This follows from Theorem 15 which shows that with cheap talk, the optimal persuasion rule can be represented as the solution of a linear program which with a number of variables and constraints which is polynomial in the primitives of the problem. The corollary follows because linear programming can be solved in polynomial time. Thus the addition of cheap talk and dynamic communication leads to a computational simplification of the persuasion problem.

Say that a dynamic persuasion (d, g) rule **contains random recommendations** if there exists $x \in X$ such that $\max\{d(x,m) : m \in M\} < 1$. In other words, there is some cheap talk claim such that conditional on that claim, the persuasion rule does not request to any piece of evidence with probability one.

Theorem 16 *1. In all persuasion problems, any optimal dynamic persuasion rule always leads to a weakly lower error probability than any optimal static persuasion rule.*

2. There exist persuasion problems such that any optimal dynamic persuasion rule leads to a strictly lower error probability than any optimal static persuasion rule.

3. Fix a persuasion problem. An optimal dynamic persuasion rule leads to a strictly lower error probability than the optimal persuasion rule only if:

(a) All optimal dynamic persuasion rules contain random recommendations.

(b) σ is not normal.[35][36]

[34] One interesting consequence of this result is that there always exists an optimal persuasion rule in which the probability that the listener accepts the speaker's request conditional on the speaker presenting the hard message that the listener requested depends only on the speaker's cheap talk claim and not on the message that the listener requested.

[35] (3a) and (3b) are not logically independent; (3a) implies (3b). I include condition (3b) in the statement of the theorem because it is of independent interest.

[36] Part 3b of Theorem 16 is similar to Theorem 5 of Bull and Watson (2007), which deals with mechanism

Before discussing the proof of the theorem, it is interesting to observe that part 3b of the Theorem 16 when combined with Corollary 3 can be used to provide an alternative proof of fact that there exists a polynomial time algorithm for finding the optimal static persuasion rule under normality.[37] Further details are given in the appendix.

For any static persuasion rule f, consider a dynamic persuasion rule (d, g) where $d(x, m) > 0$ only if $m \in \text{argmax}\{f(m) : m \in \sigma(x)\}$, and $d(x, m) > 0 \Rightarrow g(x, m) = \max\{f(m) : m \in \sigma(x)\}$. (Recall that $d(x, m) = 0 \Rightarrow g(x, m) = 0$). Then (d, g) leads to the same error probability as f, implying part 1 of the theorem.

Part 2 can be proven by example.[38] However, I give an alternative argument because it is more interesting. This argument requires one additional assumption, namely that $P \neq NP$.[39] Of course the proof by example (given in footnote 38) does not require this assumption. If $P \neq NP$, it follows from Theorem 3 that there does not exist a polynomial time algorithm that decides the question:

- Does there exist a static persuasion rule inducing an error probability of at most q?

On the other hand, it follows from Corollary 3 that there exists a polynomial time algorithm that decides the question:

design environments with evidence. One difference is in what is meant by a *static persuasion rule* in the current paper and what is meant by a *static mechanism* in Bull and Watson (2007). By a static persuasion rule, I mean a persuasion rule in which the speaker presents only a hard message $m \in \sigma(x)$ and not a cheap talk message. By a static mechanism, Bull and Watson (2007) mean a mechanism in which agents submit both a hard message and a cheap talk message. It is easy to see that in the persuasion environment, because the speaker's preferences do not depend on the state, adding cheap talk to static persuasion rules without incorporating dynamic communication would not change anything. This is not true in more general environments. For further differences between the model studied in this paper and the Bull and Watson (2007) framework, see footnote 31.

[37] In connection with this, it is also interesting to note that one can argue that the program (11) reduces to the linear programming relaxation of (6). Recall that by Corollary 2, the integer constraints in 6 are not binding.

[38] Consider the persuasion problem from Example 4, and the dynamic persuasion rule (d, g) where:

$$d(z,m) = \begin{cases} 1/2, & \text{if } z = x_i \in A, m = m_j, j \neq i; \\ 0, & \text{if } z = x_i \in A, m = m_i; \\ 1, & \text{if } z = y_i \in R, m = m_i; \\ 0, & \text{if } z = y_i \in R, m = m_j, j \neq i; \end{cases} \qquad g(z,m) = \begin{cases} 1, & \text{if } z = x_i \in A, m = m_j, j \neq i; \\ 0, & \text{if } z = x_i \in A, m = m_i; \\ 1/2, & \text{if } z = y_i \in R, m = m_i; \\ 0, & \text{if } z = y_i \in R, m = m_j, j \neq i; \end{cases}$$

(d, g) induces an error probability of $3/2p_y$ whereas the analysis in Example 4 established that the optimal static persuasion rule achieves an error probability of $2p_y$ in this problem. This example establishes part 2 of the theorem.

It is interesting to observe that the error probability achievable by the optimal dynamic persuasion rule is the same as the value of the program (2)-(4) if we relax the integer constraints (i.e., replace the constraints (4) by the constraints (5)). This might lead one to conjecture that in general one can find the error probability at the optimal dynamic persuasion rule by solving the linear programming relaxation of the integer program (1). However this conjecture is false; while the linear programming relaxation of (1) always gives a lower bound on the error probability at the optimal dynamic rule, sometimes the optimal dynamic rule fails to achieve this lower bound.

[39] This is a famous unproven theorem. If $P = NP$ then large parts of theoretical computer science would undermined. For a discussion of the P vs. NP problem, see Papadimitriou (1994).

- Does there exist a dynamic persuasion rule inducing an error probability of at most q?

If optimal static persuasion rules always led to the same error probability as optimal dynamic persuasion rules, then the polynomial time algorithm deciding the second question could also be used to decide the first question, a contradiction. This argument is interesting because it relates the fact that cheap talk and dynamic communication simplify the computation of the optimal persuasion rule to the fact that cheap talk and dynamic communication are valuable to the listener in terms of achieving his objective.[40]

Part 3a of Theorem 16 follows from the fact that any dynamic persuasion rule (d, g) without random recommendations is equivalent to the static persuasion rule f with $f(m) := \max\{g(x, m) : x \in X\}$. To prove part 3b of the theorem, assume that the persuasion problem is normal, and consider any incentive-compatible dynamic persuasion rule (d, g)–in other words, consider any dynamic persuasion rule which gives the speaker the incentive to make a truthful cheap talk claim (of course, once the speaker does this, he will also have an incentive to obey the listener's request). Now construct the static persuasion rule f such that $f(m_x) = \sum_{m\in\sigma(x)} d(x, m)g(x, m)$. Observe that Definition 1 and incentive compatibility of (d, g) implies that if $m_x = m_y$, then $\sum_{m\in\sigma(x)} d(x, m)g(x, m) = \sum_{m\in\sigma(y)} d(y, m)g(y, m)$, so f is well defined. So at x, if the speaker sends m_x, he will be accepted by f with the same probability as he would have been accepted by (d, g). It cannot be that the at x, the speaker has a message m which would lead to acceptance with a higher probability. If he did, then $m = m_y$ for some $y \in X$, and by normality $\sigma(y) \subseteq \sigma(x)$, so incentive compatibility of (d, g) implies that $f(m_y) = \sum_{m\in\sigma(y)} d(y, m)g(y, m) \leq \sum_{m\in\sigma(x)} d(x, m)g(x, m) = f(m_x)$, a contradiction. This establishes 3b.

9 Conclusion

This paper has established that the complexity of the problem of finding optimal rules of persuasion depends on qualitative properties of the evidence, and in particular *normality*. The paper presents an algorithm for finding the optimal rule under normality. It would be interesting to extend the analysis to situations with multiple speakers, private information on the part of the listener, as well as situations in which messages have costs which depend on the state. A more general goal would be to extend the methodology of this paper to a more general analysis of mechanism design with evidence and limitations on communication. The problem of determining how to respond to messages when the availability of messages

[40] A similar argument can be used to show that if $P \neq NP$, there exists a normal form game Γ and a player i in Γ such that player i's utility in the best correlated equilibrium of Γ is strictly greater than i's utility in the best Nash equilibrium of Γ. Of course this can also be proven without assuming $P \neq NP$, simply by constructing an example.

depends on the state is inherently combinatorial, and therefore it is important to explicitly account for the computational costs associated with such optimization.

Appendix: Proofs

Proof of Theorem 2

For $Z = A, R$, let $M_Z := \{m \in M : \forall x \in X, m \in \sigma(x) \Rightarrow x \in Z\}$. Let $f'(m) := f(m)$ for all $m \in M \setminus (M_A \cup M_R)$, $f'(m) := 1$ for all $m \in M_A$ and $f'(m) := 0$ for all $m \in M_R$. If $\alpha(f) \neq \alpha(f')$, then the error probability of f' is lower than that of f, a contradiction. So $\alpha(f) = \alpha(f')$, implying f' is optimal. Since ζ is a best reply to f, for $m \in M_A$, $\sum_{x:m\in\sigma(x)} \zeta(x,m)p_x > 0 \Rightarrow f(m) = 1$, and for $m \in M_R$, $\sum_{x:m\in\sigma(x)} \zeta(x,m)p_x > 0 \Rightarrow f(m) = 0$, implying that f' is a best reply to ζ, and since $\alpha(f) = \alpha(f')$, ζ is a best reply to f'. For $m \in M$ and $Z = A, R$, define $P(m,Z) := \sum_{x\in Z:m\in\sigma(x)} p_x$. For $x \in X, m \in \sigma(x)$, and small $\epsilon > 0$, define:

$$\zeta^\epsilon(x,m) = \begin{cases} \zeta(x,m) - \gamma_x^\epsilon, \text{ if } \zeta(x,m) > 0; \\ \epsilon, \text{if } \zeta(x,m) = 0 \text{ and either } (x \in A \text{ and } f'(m) = 1) \text{ or } (x \in R \text{ and } f'(m) = 0); \\ \epsilon^2, \text{if } \zeta(x,m) = 0 \text{ and either } (x \in A \text{ and } f'(m) = 0) \text{ or } (x \in R \text{ and } f'(m) = 1); \\ \epsilon/P(m,Z), \text{ if } \zeta(x,m) = 0,\ f'(m) \in (0,1), \text{ and } x \in Z \text{ for } Z = A, R. \end{cases}$$

where for each $x \in X$, γ_x^ϵ is chosen so that $\sum_{m\in\sigma(x)} \zeta^\epsilon(x,m) = 1$. $\zeta^\epsilon \to \zeta$ as $\epsilon \to 0$. Next observe that for $m \in M$ with $\sum_{x\in X:m\in\sigma(x)} \zeta(x,m)p_x = 0$, we have:

$$\frac{\sum_{x\in A:m\in\sigma(x)} \zeta^\epsilon(x,m)p_x}{\sum_{x\in X:m\in\sigma(x)} \zeta^\epsilon(x,m)p_x} \to \begin{cases} 1, & \text{if } f'(m) = 1 \\ 1/2, & \text{if } f'(m) \in (0,1) \\ 0, & \text{if } f'(m) = 0 \end{cases} \quad \text{as } \epsilon \to 0. \tag{13}$$

(13) depends on the fact that for all $m \in M_A$, $f'(m) = 1$ and for all $m \in M_R$, $f'(m) = 0$. This establishes that (ζ, f') is a sequential equilibrium. □

Proof of Theorem 3

Let U be a set and $\mathcal{S}$ a family of subsets of U such that $\bigcup \mathcal{S} = U$. The SET COVER PROBLEM is the problem of finding a minimal cardinality subset $\mathcal{T}$ of $\mathcal{S}$ such that $\bigcup \mathcal{T} = U$. This problem is known to be NP-hard. I will prove that the persuasion problem (without normality) is NP-hard by reduction from the SET COVER PROBLEM. Consider an instance of the SET COVER PROBLEM $(U, \mathcal{S})$. For each $S \in \mathcal{S}$ construct a state x_S, and let $X = U \cup \{x_S : S \in \mathcal{S}\}$. Let $A = U$ and $R = \{x_S : S \in \mathcal{S}\}$. For each $x \in A$, let $p_x := \frac{2|R|}{(2|A|+1)|R|}$ and for each $y \in R$, let $p_y := \frac{1}{(2|A|+1)|R|}$ For each $S \in \mathcal{S}$, let there be a message m_S, and define σ so that $m_S \in \sigma(x) \Leftrightarrow x \in S \cup \{x_S\}$. We know that there is an optimal deterministic persuasion rule, and any optimal deterministic persuasion rule can be associated with the set K of messages that it accepts. Given the specification of the above persuasion problem, it is clear that any optimal persuasion rule must accept all states in $A = U$, or more precisely, must accept some message available at x for all $x \in A$. The goal is then to do this while accepting the minimum number of states in R, which amounts to finding a minimal collection $K \subseteq M$ such that for all $x \in A$, $K \cap \sigma(x) \neq \emptyset$. This is equivalent to finding a minimal cardinality subset $\mathcal{T}$ of $\mathcal{S}$ that covers U (i.e., $\bigcup \mathcal{T} = U$), which is the SET COVER PROBLEM. □

Proof of Theorem 7

Fix a flow φ. Define $\mu_x^* = 1$ if $x \in [(V^\varphi \cap X) \setminus A] \cup [V^\varphi \cap R]$ and $\mu_x^* = 0$ otherwise. The first step is to verify that (*) if φ is a maximum flow in the persuasion problem, then the vector $\{\mu_x^*\}_{x \in X}$ is a solution to (6). Let $\nu_x = 1 - \mu_x$ if $x \in A$ and $\nu_x = \mu_x$ if $x \in R$, $w_x = p_x$ if $x \in A$ and $w_x = -p_x$ if $x \in R$. Then (6) is equivalent to: $\max_{\{\nu_x\}_{x\in X}} \sum_{x \in X} w_x \nu_x$ s.t. $\nu_x \in \{0, 1\}$ for all $x \in X$ and $\nu_x \leq \nu_y$ for all $x \in A$ and $y \in R$ such that $\sigma(x) \subseteq \sigma(y)$. Given this equivalence, (*) now follows from the construction in Sections 4 and 5 of Picard (1976). A clear explanation of this construction can be found in Section 2 of Hochbaum (2001) or Section 2 of Hochbaum (2004).

Let f be as in (7) and ζ as in (9). (That ζ indeed defines a probability distribution over messages at each state follows from the flow conservation constraints). I now argue that: (**) If $\sigma(x) \cap \{m : f(m) = 1\} \neq \emptyset$, then then ζ puts probability 1 on messages m such that $f(m) = 1$ at x. In each of four cases, I cite the facts that imply (**). Case 1 $x \in A \cap V^\varphi$: immediate. Case 2 $x \in R \cap V^\varphi$: in this case $\varphi(x, t) = p_x > 0$. For all $y \in A$, if $\varphi(y, x) > 0$, then $y \in V^\varphi$. Case 3 $x \in R \setminus V^\varphi$: for all $y \in A$, if $(y, x) \in E$ (i.e., $m_y \in \sigma(x)$), then $y \notin V^\varphi$, so $\sigma(x) \cap \{m : f(m) = 1\} = \emptyset$. Case 4 $x \in A \setminus V^\varphi$: in this case $\varphi(s, x) = p_x$ which is only possible if there exists $y \in R$ such that $m_y \in \sigma(x)$ and $\varphi(x, y) > 0$. Assume for contradiction that there exists $z \in A \cap V^\varphi$ such that $m_z \in \sigma(x)$. By normality, $m_z \in \sigma(y)$, implying that $y \in V^\varphi$, and since $\varphi(x, y) > 0$, $x \in V^\varphi$, a contradiction. This establishes (**). The considerations in this paragraph also establish that $\alpha(f, x) = 1$ if and only if $x \in [(V^\varphi \cap X) \setminus A] \cup [V^\varphi \cap R]$; the first paragraph now implies that f is an optimal rule. V^φ is a minimum cut (see Section 8.1 of Korte and Vygen (2006)), and it follows from the max-flow min-cut theorem that value$(\varphi) = \sum_{x \in A \setminus V^\varphi} p_x + \sum_{x \in R \cap V^\varphi} p_x$, which implies that value(φ) is the error probability.

If $f(m) = 1$, then $m = m_x$ for some $x \in A \cap V^\varphi$. We have:

$$p_x \geq \varphi(s, x) = \sum_{y \in R : m_x \in \sigma(y)} \varphi(x, y) = \sum_{y \in R : m_x \in \sigma(y)} \frac{\varphi(x, y)}{\varphi(y, t)} p_y = \sum_{y \in R : m_x \in \sigma(y)} \zeta(y, m_x) p_y,$$

where the inequality follows from the capacity constraints, the first equality from the flow conservation constraints, and the second equality from the fact that since $x \in A \cap V^{\varphi}$, then $p_x = \varphi(s,x)$. So in equilibrium, conditional on receiving m, the listener prefers to accept m. Next, observe that if $m = m_x$ and $f(m) = 0$ for some $x \in A$, then $x \in A \setminus V^{\varphi}$, implying that $p_x = \varphi(s,x)$. So:

$$p_x = \varphi(s,x) = \sum_{y \in R: m_x \in \sigma(y)} \varphi(x,y) \leq \sum_{y \in R: m_x \in \sigma(y), \varphi(y,t) \neq 0} \frac{\varphi(x,y)}{\varphi(y,t)} p_y \leq \sum_{y \in R: m_x \in \sigma(y)} \zeta(y, m_x) p_y,$$

where the second equality follows from the flow conservation constraints and the inequality follows the capacity constraints of the form $\varphi(y,t) \leq p_y$, and the fact that $\varphi(y,t) = 0 \Rightarrow \varphi(z,y) = 0$ for all $(z,y) \in E$. The last inequality follows from the definition of ζ, and is not necessarily an equality because the term preceding the inequality does not include the probability that m is sent in a state $y \in R$ with $\varphi(y,t) = 0$. It follows that in equilibrium, conditional on seeing m, the listener prefers to reject m. Finally for any m with $f(m) = 0$ and $m \neq m_x$ for all $x \in A$, if m is sent at all, it is only sent in a state in R, so again the listener prefers to reject. It follows that (ζ, f) is an equilibrium. □

Proof of Theorem 8

Let Π be the set of all symmetries. Let f be any optimal persuasion rule. Then for every $(\pi, \xi) \in \Pi$, $f \circ \xi$ is an optimal rule. Define $f^*(m) := \max\{f \circ \xi(m) : (\pi, \xi) \in \Pi\}$. Then for all x, $\alpha(f^*, x) = \max\{\alpha(f \circ \xi) : (\pi, \xi) \in \Pi\}$. It follows from (*) in the proof of Theorem 10 that f^* is also an optimal rule. As a notational matter, let $(\pi, \xi) \circ (\pi', \xi') := (\pi \circ \pi', \xi \circ \xi')$, where $\circ$ denotes the composition of functions. For for any $(\pi', \xi') \in \Pi$ and any $m \in M$: $f^* \circ \xi'(m) = \max\{f \circ \xi(\xi'(m)) : (\pi, \xi) \in \Pi\} = \max\{f \circ \xi''(m) : (\pi'', \xi'') \in \Pi \circ (\pi', \xi')\} = \max\{f \circ \xi''(m) : (\pi'', \xi'') \in \Pi\} = f^*(m)$. The third equality follows from the fact that Π is a group under $\circ$. So f^* is symmetric. □

Proof of Theorem 9

Assume f is a symmetric optimal rule which does not reject some message m: $f(m) > 0$. By optimality, there exists x^* with $\sum_i x_i^* =: k \geq \ell$ and $J \subseteq \{1, \ldots, n\}$, $|J| \leq h$ with $f(x_J^*) > 0$. In other words, optimality implies that if f does not reject all messages, there must be some message m that f accepts with positive probability which is available at some state in A. For every q with $h \leq q < k$, there exists $y^q \in X$ with $\sum_i y_i^q = q$ and $y_J^q = x_J^*$. By symmetry, f accepts some message with positive probability from every state with $h \leq \sum_i x_i \leq k$. It follows that the optimal rule f accepts every message of the form $\{(1,j) : j \in J\}$ with $|J| = h$ with positive probability, and by symmetry it must assign all these messages the same probability, in fact 1 (since otherwise it would be optimal to reject all these messages as well). Note that any optimal rule must reject any message which reveals that any component is 0 or shows less than h components, because in this way it is possible to reject every type who has at most h 1's without reducing the probability of acceptance conditional on A. But now consider the rule f' which agrees with f except on message $\{(1,j) : 1 \le j \le h\}$, which it rejects. Such a rule would reject on an additional state in R, without rejecting on any additional states in A, and hence would do better, contradiction. $\square$

Proof of Theorem 10

Assume normality. Then $P := \{\alpha(f) : f \in F\} = \{\beta \in [0,1]^X : \forall x, y \in X, \sigma(x) \subseteq \sigma(y)\}$ and the persuasion problem is equivalent to $\max \sum_{x \in A} \beta_x p_x - \sum_{y \in R} \beta_y p_y$ s.t. $\beta \in P$. The objective is modular on $[0,1]^X$, hence also supermodular, and P is a subcomplete sublattice of $[0,1]^X$ under the product order. Corollary 2.7.1 of Topkis (1998) implies that there exist most and least difficult optimal rules. In fact (*) the subset of optima in P forms a sublattice of P and hence $[0,1]^X$. Since the extreme points of P are integral, the most and least difficult optimal rules are deterministic. The comparative statics results follow from Theorem 2.8.2 of Topkis (1998). See also Section 3.7.2 and in particular Theorem 3.7.4 of Topkis (1998) which analyze the comparative statics of the maximal closure problem and presents a similar analysis.

For a counter-examples to the comparative statics results without normality, assume $X = \{0,1,2\}, M = \{m_1, m_2\}, \sigma(0) = \{m_1, m_2\}, \sigma(1) = \{m_1\}, \sigma(2) = \{m_2\}, p_0 = 1/2, p_1 = 3/10, p_2 = 1/5, A = \{0\}$. The message structure is not normal. The unique optimal rule accepts m_2 and rejects m_1. Shifting 2/5 probability mass from $1 \in R$ to $0 \in A$ will create a situation where the unique optimal rule accepts m_1 and rejects m_2 which is neither more nor less difficult than the optimal rule in the original problem. Go back to the probabilities in the original problem and move state 1 from R to A. Then the optimal rule again accepts m_1 and rejects m_2. Finally reset the probabilities so that $p_0 > p_1 = p_2$. Then there are two optimal rules: (i) accept m_1, reject m_2, (ii) accept m_2, reject m_1, so there exists neither a most nor least difficult optimal rule.

Proof of Theorem 14

Start with a persuasion problem $Q := (X, M, \sigma, A, p)$. First enumerate the states in A so that $A = \{x_1, x_2, \ldots, x_n\}$. For all messages $m \in \sigma(x)$, iteratively ask the best reply oracle whether x_1 sends m in a best reply to the optimal rule, until the oracle says yes at some message m_1. Then for each $m \in \sigma(x_2)$, ask whether there exists a best reply to the optimal rule in which x_1 sends m_1 and x_2 sends m until you find a message m_2 for which the answer is yes. Iterate until you arrive at a sequence of messages $\{m_1, m_2, \ldots, m_n\}$ such that $m_i \in \sigma(x_i)$ and there exists an optimal rule at which x_1 sends m_1, x_2 sends m_2, and $\ldots$ x_n sends m_n. Using a Turing machine with an oracle (See Definition 14.3 on p.

339 of Papadimitriou (1994)), each step requires writing a state x and a message m on the query string for the oracle, and there are at most $|X| \cdot |M|$ such steps, so this can be done in polynomial time.[41] Next consider the persuasion problem $Q^* := (X, M, \sigma^*, A, p)$, which agrees with Q except on the message correspondence σ^* which is defined so that $\sigma^*(x_i) = \{m_i\}$ for all $x_i \in A$, and $\sigma^*(y) = \sigma(y)$ for all $y \in R$.

Lemma 1 *Theorem 7 is still true if the assumption of normality is weakened to normality on A; the corresponding maximal flow problem mentioned in the theorem is constructed in exactly the same way.*

Proof. Suppose that a persuasion problem $O = (X, M, \sigma, A, p)$ is normal on A. Then create a new persuasion problem $O' = (X, M \cup \{m^y : y \in Y\}, \sigma', A, p)$ where σ agrees with σ' on A, and $\sigma(y) = \sigma(y) \cup \{m^y\}$ for some new message m^y. Then O' is normal. Moreover the network corresponding to O is the same as the network corresponding to O'. (Notice in particular that for all $x \in A$ and $y \in R$, $\sigma(x) \subseteq \sigma(y) \Leftrightarrow \sigma'(x) \subseteq \sigma'(y)$, and these inclusion relations determine the edges of the network). Clearly adding the messages m^y for $y \in R$ will not alter the set of optimal rules in any essential way because the listener has no incentive to accept messages which can only be used at R. Moreover, since the networks corresponding to the two persuasion problems are the same, the credible implementations for the two persuasion problems will be essentially the same. $\square$

Observe that Q^* is trivially normal on A since in every state $x \in A$, only one message is available. It follows from Lemma 1 that there exists a polynomial time algorithm finding an optimal persuasion rule f^{**} and a credible implementation (ζ^{**}, f^{**}) in Q^*. Let f be any persuasion rule. At every state in R, f induces the same acceptance probability in Q^* as in Q. In every state in A, f induces a weakly lower acceptance probability in problem Q^* than in Q. It follows that every persuasion rule induces a weakly higher error probability in Q^* than in Q. Let f^* be the persuasion rule which is optimal in Q such that it is a best reply in x_1 to send m_1, at x_2 to send $m_2, \ldots$, and at x_n to send m_n. f^* induces the same error probability in Q and Q^* and hence is optimal in Q^*. Since f^{**} is also optimal in Q^*, and because all rules induce weakly lower error probability in Q than in Q^*, it follows that f^{**} is also an optimal persuasion rule in Q. We can think of ζ^{**} as a speaker strategy also in Q where $\zeta^{**}(x_i, m) = 0$ for $x_i \in A$ and $m \in \sigma(x_i) \setminus \{m_i\}$. Finally notice that ζ^{**} is still a best reply to f^{**} in Q (because the speaker's choice set has does not change in R when moving from Q^* to Q). So if ζ^{**} is not a best reply to f^{**}, that means that there is a state $x \in A$ at

[41] Formally, what is the language that the oracle can decide? The oracle can decide the language which consists of a description of the persuasion problem Q followed by a list of states x_i in A and messages $m_i \in \sigma(x_i)$ such that there is a best reply to some optimal rule f^* in the persuasion problem Q in which the speaker sends m_1 in x_1, m_2 in x_2, $\ldots$, and m_n in x_n. Of course, the Turing machine must copy the description of Q on the query tape for the oracle. However, it is important to notice, that in the course of finding an optimal rule for Q, the oracle need not be asked any questions about any persuasion problem other than Q.

which the speaker can increase its acceptance probability by switching its strategy and no state $y \in R$ that can do this, which implies that the error probability induced by f^{**}–which is calculated assuming the speaker best replies–is lower in Q than in Q^*, a contradiction. $\square$

Proof of Theorem 15

Consider any incentive compatible persuasion rule (d', g'). In other words, (d', g') gives the speaker the incentive to truthfully report the state and then to show the listener whichever message he requests. Defining $\beta(x, m) := d'(x, m)g'(x, m)$, at state x, the speaker's utility to playing straightforwardly is: $\sum_{m \in \sigma(x)} \beta(x, m)$. Notice that if the speaker misrepresents the state to be y when the state is x at the initial stage, then at the second stage the speaker will have the incentive to present the hard message that the listener requests if that message is available, so that his utility will be $\sum_{m \in \sigma(x) \cap \sigma(y)} \beta(y, m)$. These observations show that the constraints in (11) correspond to the speaker's incentive constraints (as well as the constraint that the ultimate probability of acceptance conditional on any cheap talk message must be between zero and one). Next note that if (d, g) is defined as in (12) with respect to $\beta(x, m) := d'(x, m)g'(x, m)$, then the speaker's payoff to any behavior which is optimal conditional on his cheap talk report is unchanged. The proof is completed by observing that the objective corresponds to the probability of error. $\square$

Alternative Proof that Under Normality, Static Persuasion Problem is Polynomial Using Corollary 3 and Theorem 16

Consider a normal persuasion problem, and consider an optimal dynamic rule (d, g). The proof of Part 3b of Theorem 16 established that the static persuasion rule f such that $f(m_x) := \sum_{m \in \sigma(x)} d(x, m)g(x, m)$ and $f(m) = 0$ if $m \neq m_x$ for all x is well defined, and is moreover an optimal static persuasion rule. Corollary 3 implies that (d, g) can be found in polynomial time, which implies that f can also be found in polynomial time. $\square$

References

Ben-Porath, E., and B. Lipman (2008): "Implementation and Partial Provability," Unpublished Manuscript.

Bull, J., and J. Watson (2004): "Evidence Disclosure and Verifiability," *Journal of Economic Theory*, 118.

——— (2007): "Hard Evidence and Mechanism Design," *Games and Economic Behavior*, 58, 75–93.

Conitzer, V., and T. Sandholm (2008): "New Complexity Results adout Nash Equilibria," *Games and Economic Behavior*, 63, 621–641.

Dasgupta, S., C. Papadimitriou, and U. Vazirani (2008): *Algorithms*. McGraw Hill.

Deneckere, R., and S. Severenov (2008): "Mechanism Design with Partial State Verifiability," *Games and Economic Behavior*, 64, 487–513.

Deneckere, R., and S. Severinov (2001): "Mechanism Design and Communication Costs," Unpublished Manuscript, University of Wisconsin.

Dziuda, W. (2007): "Strategic Arumentation," Unpublished Manuscript.

Edmonds, J., and R. Karp (1972): "Theoretical Improvements in Algorithmic Efficiency for Network Flow Problems," *Journal of the ACM*, 19, 248–264.

Fishman, M., and K. Hagerty (1990): "The Optimal Amount of Discretion to Allow in Disclosure," *The Quarterly Journal of Economics*, 105, 427–444.

Ford, L., and D. Fulkerson (1957): "A Simple Algorithm for Finding Maximal Network Flows and an Application to the Hitchcock Problem," *Canadian Journal of Mathematics*, 9, 399–404.

Forges, F., and F. Koessler (2005): "Communication Equilibria with Partially Verifiable Types," *Journal of Mathematical Economics*, 41, 793–811.

Garey, M., and D. Johnson (1979): *Computers and Intractibility*. W. H. Freeman and Company.

Gilboa, I., and E. Zemel (1989): "Nash and Correlated Equilibria: Some Complexity Considerations," *Games and Economic Behavior*, 1, 80–93.

Glazer, J., and A. Rubinstein (2001): "Debates and Decisions: On a Rationale of Argumentation Rules," *Games and Economic Behavior*, 36, 158–173.

——— (2004): "On the Optimal Rules of Persuasion," *Econometrica*, 72, 1715–1736.

——— (2006): "A Study in the Pragmatics of Persuasion: A Game Theoretical Approach," *Theoretical Economics*, 1, 395–410.

Green, J., and J. Laffont (1986): "Partially Verifiable Information and Mechanism Design," *Review of Economic Studies*, 53, 447–456.

Green, J., and J.-J. Laffont (1977): "Characterization of Satisfactory Mechanisms for the Revelation of Preferences for Public Goods," *Econometrica*, 45, 427–438.

GROSSMAN, S. (1981): “The Informational Role of Warranties and Private Disclosure about Product Quality,” *Journal of Law and Economics*, 24, 461–483.

GROSSMAN, S., AND O. HART (1980): “Disclosure Laws and Takeover Bids,” *Journal of Finance*, 35, 323–334.

HOCHBAUM, D. (2001): “A New-Old Algorithm for Minimum-cut and Maximum-flow in Closure Graphs,” *Networks*, 37, 171–193.

——— (2004): “Selection, Provisioning, Shared Fixed Costs, Maximum Closure, and Implications on Algorithmic Methods Today,” *Management Science*, 50, 709–723.

——— (2008): “Network Flow and Graph Algorithms,” Lecture Notes, IEOR 266 Fall 2008, Department of Industrial Engineering and Operations Research, University of California Berkeley.

KARTIK, N., AND O. TERCIEUX (2009): “Implementation with Evidence: Complete Information,” Unpublished Manuscript.

KORTE, B., AND J. VYGEN (2006): *Combinatorial Optimization*. Springer, 3 edn.

LIPMAN, B., AND D. SEPPI (1995): “Robust Inference in Communication Games with Partial Provability,” *Journal of Economic Theory*, 66, 370–405.

MILGROM, P. (1981): “Good News and Bad News: Representation Theorems and Applications,” *The Bell Journal of Economics*, 2, 380–391.

MILGROM, P., AND J. ROBERTS (1986): “Relying on Information of Interested Parties,” *Rand Journal of Economics*, 17, 18–32.

MYERSON, R. (1986): “Multistage Games with Communication,” *Econometrica*, 54, 328–358.

PAPADIMITRIOU, C. (1994): *Computational Complexity*. Addison-Wesley Publishing Company.

PICARD, J.-C. (1976): “Maximal Closure of a Graph and Applications to Combinatorial Problems,” *Management Science*, 22, 1268–1272.

SHER, I. (2008): “Credibility and Determinism in a Game of Persuasion,” Unpublished Manuscript.

SHIN, H. (1994a): “The Burden of Proof in a Game of Persuasion,” *Journal of Economic Theory*, 64, 253–264.

——— (1994b): "News Management and the Value of Firms," *Rand Journal of Economics*, 25, 58–71.

——— (2003): "Disclosure and Asset Returns," *Econometrica*, 71, 105–133.

Singh, N., and D. Wittman (2001): "Implementation with Partial Verification," *Review of Economic Design*, 6, 63–84.

Sipser, M. (2006): *Intorduction to the Theory of Computation.* Thomson Course Technology, 2nd edn.

Topkis, D. (1998): *Supermodularity and Complementarity.* Princeton University Press.